UNIVERSIDADE FEDERAL DO CEARÁ

CENTRO DE CIÊNCIAS E TECNOLOGIA

DEPARTAMENTO DE ENGENHARIA DE TELEINFORMÁTICA

PROGRAMA DE PÓS-GRADUAÇÃO EM TELEINFORMÁTICA

MESTRE EM ENGENHARIA DE TELEINFORMÁTICA

CARLOS ALFREDO CORDEIRO DE VASCONCELOS FILHO

IQAEVOLNET: A NOVEL UNSUPERVISED EVOLUTIONARY IMAGE ENHANCEMENT ALGORITHM ON CHEST X-RAY SCANS

FORTALEZA

2024

CARLOS ALFREDO CORDEIRO DE VASCONCELOS FILHO

IQAEVOLNET: A NOVEL UNSUPERVISED EVOLUTIONARY IMAGE ENHANCEMENT ALGORITHM ON CHEST X-RAY SCANS

Dissertation submitted to the Programa de Pós-graduação em Teleinformática of the Centro de Ciências e Tecnologia of the Universidade Federal do Ceará, as a partial requirement for obtaining the title of Master in Engenharia de Teleinformática. Concentration Area: Engenharia de Teleinformática

Advisor: Prof. Dr. Victor Hugo Costa de Albuquerque

Co-advisor: Prof. Dr. Paulo Cesar Cortez

FORTALEZA

2024

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Sistema de Bibliotecas
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

F498i Filho, Carlos Alfredo Cordeiro de Vasconcelos.
IQAEvolNet: A Novel Unsupervised Evolutionary Image Enhancement Algorithm on Chest X-Ray Scans / Carlos Alfredo Cordeiro de Vasconcelos Filho. – 2024.
78 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia de Teleinformática, Fortaleza, 2024.
Orientação: Prof. Dr. Victor Hugo Costa de Albuquerque.
Coorientação: Prof. Dr. Paulo Cesar Cortez.

1. Aprimoramento de imagens. 2. Rede neural convolucional. 3. Aprendizado não supervisionado. 4. Algoritmo genético. 5. Raio-X. I. Título.

CDD 621.38

CARLOS ALFREDO CORDEIRO DE VASCONCELOS FILHO

IQAEVOLNET: A NOVEL UNSUPERVISED EVOLUTIONARY IMAGE ENHANCEMENT ALGORITHM ON CHEST X-RAY SCANS

Dissertation submitted to the Programa de Pós-graduação em Teleinformática of the Centro de Ciências e Tecnologia of the Universidade Federal do Ceará, as a partial requirement for obtaining the title of Master in Engenharia de Teleinformática. Concentration Area: Engenharia de Teleinformática

Approved on:

EXAMINATION BOARD

Prof. Dr. Victor Hugo Costa de Albuquerque (Advisor)
Universidade Federal do Ceará (UFC)

Prof. Dr. Paulo Cesar Cortez (Co-advisor)
Universidade Federal do Ceará (UFC)

Prof. Dr. Francisco Nauber Bernardo Gois
Controladoria Geral do Estado (CGE)

Prof. Dr. Renato Willian Rodrigues de Souza
Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE)

À minha família, por sua capacidade de acreditar em mim e investir em mim. Mãe, seu cuidado e dedicação foi que deram, em alguns momentos, a esperança para seguir. Pai, sua presença significou segurança e certeza de que não estou sozinho nessa caminhada.

ACKNOWLEDGEMENTS

Ao Prof. Dr. Victor Hugo C. de Albuquerque por me orientar em minha dissertação de mestrado.

Ao Prof. Dr. Paulo Cesar Cortez por me auxiliar em meus estudos.

Aos meus pais, irmãs e sobrinhos, que nos momentos de minha ausência dedicados ao estudo superior, sempre fizeram entender que o futuro é feito a partir da constante dedicação no presente!

Agradeço a todos os professores por me proporcionar o conhecimento não apenas racional, mas a manifestação do caráter e afetividade da educação no processo de formação profissional, por tanto que se dedicaram a mim, não somente por terem me ensinado, mas por terem me feito aprender.

E à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) pelo financiamento da pesquisa de mestrado via bolsa de estudos.

"O sonho é que leva a gente para frente. Se a gente for seguir a razão, fica aquietado, acomodado."

(Ariano Suassuna)

ABSTRACT

X-ray images are widely used in the medical field due to their low cost and non-invasive nature, but they suffer from noise problems related to equipment or environmental factors. There are multiple solutions in the literature to combat this problem, with the main one being the use of non-learning image processing algorithms for X-ray image enhancement. In other areas of application, such as low-light and underwater images, there is an extensive use of artificial intelligence models for image enhancement tasks, but the training of artificial intelligence models for medical image enhancement encounters significant challenges. In supervised learning, obtaining a dataset with authentic noisy images and their manually enhanced counterparts as labels is imperative. When dealing with medical images it can be difficult to have access to high-quality/low-quality pairs because of the restrictive context where these images are taken. To deal with this problem, this paper introduces an innovative approach to unsupervised learning for chest x-ray image enhancement. The suggested approach begins with the pre-training of a model using multiple image enhancement algorithms as reference to establish an initial set of solutions. Following this, an evolutionary algorithm is employed to refine these initial solutions. It incorporates two image enhancement metrics, Entropy and the Natural Image Quality Evaluator(NIQE), along with Structural Similarity Index as fitness indicators. We tested our method in a Chest X-ray dataset and our findings demonstrate that our method achieved a better NIQE, 4.05 compared to 4.24, and a faster processing time, 2.95 milliseconds compared to 0.195 seconds, in relation to the state-of-the-art algorithm with the best NIQE and entropy. We showed that our algorithm outperforms state-of-the-art algorithms in NIQE and processing time.

Keywords: Image enhancement. Convolutional Neural Network. Unsupervised learning. Evolutionary algorithms. Medical image. X-ray.

RESUMO

As imagens de raio-X são amplamente utilizadas no campo médico devido ao seu baixo custo e natureza não invasiva, mas sofrem de problemas de ruído relacionados ao equipamento ou a fatores ambientais. Existem várias soluções na literatura para combater esse problema, sendo a principal o uso de algoritmos de processamento de imagem sem aprendizado para o aprimoramento de imagens de raio-X. Em outras áreas de aplicação, como imagens com pouca luz e subaquáticas, há um amplo uso de modelos de inteligência artificial para tarefas de aprimoramento de imagem, mas o treinamento de modelos de inteligência artificial para o aprimoramento de imagens médicas enfrenta desafios significativos. No aprendizado supervisionado, a obtenção de um conjunto de dados com imagens ruidosas e suas contrapartes aprimoradas é imperativa. Ao lidar com imagens médicas, pode ser difícil ter acesso a pares de alta qualidade/baixa qualidade devido ao contexto restritivo em que essas imagens são capturadas. Para lidar com esse problema, este artigo apresenta um algoritmo inovador de aprendizado não supervisionado para o aprimoramento de imagens de raio-X de tórax. A abordagem sugerida começa com o pré-treinamento de um modelo usando vários algoritmos de aprimoramento de imagem como referência para estabelecer um conjunto inicial de soluções. Em seguida, um algoritmo evolutivo é empregado para refinar essas soluções iniciais. Esse algoritmo incorpora duas métricas de aprimoramento de imagem, Entropia e o Natural Image Quality Evaluator(NIQE), juntamente com o Índice de Similaridade Estrutural como indicadores de aptidão. Nosso método foi testado em um conjunto de dados de Raio-X de tórax e nossos resultados demonstram que nossa abordagem alcançou uma pontuação NIQE melhor de 4.05 em comparação com 4.24, e um tempo de processamento mais rápido de 2.95 milissegundos em comparação com 0.195 segundos, em relação ao algoritmo estado-da-arte com as melhores pontuações de NIQE e entropia. Mostramos que nosso algoritmo supera os algoritmos estado-da-arte em termos de pontuação NIQE e tempo de processamento.

Palavras-chave: Aprimoramento de imagens. Rede neural convolucional. Aprendizado não supervisionado. Algoritmo genético. Imagem médica. Raio-X.

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

CT	Computed tomography
MRI	Magnetic resonance imaging
CNN	Convolutional neural network
GAN	Generative adversarial network
IQA	Image quality assessment
$NIQE$	Natural Image Quality Evaluator
$SSIM$	Structural Similarity Index Measure
UM	Unsharp Masking
HEF	High-Frequency Emphasis Filtering
$CLAHE$	Contrast Limited Adaptive Histogram Equalization
$ATACE$	Automatic Tissue Attenuation-based Contrast Enhancement
$ATACE$	Triple Clipped Dynamic Histogram Equalization
$*$	Convolutional operator
$ReLU$	Rectified Linear Unit
$tanh$	Hyperbolic tangent
SGD	Stochastic gradient descent
HE	Histogram equalization
$SSCL$	Self-Similarity Contrastive Learning
MR	Magnetic resonance
$PSNR$	Peak Signal-to-Noise Ratio
$BRISQUE$	Blind/Referenceless Image Spatial Quality Evaluator
$PIQE$	Perceptual Image Quality Evaluator
FFT	Fast Fourier transform
$IFFT$	Inverse fast Fourier transform
$h(r_k)$	Histogram
$cdf(r_k)$	Cumulative distribution function

exp	Natural exponential function
log	Natural logarithmic function
μ	Mean
σ	Standard deviation
PDF	Probability distribution function
DWT	Discrete Wavelet Transform
SVD	Singular value decomposition
SF	Spatial frequency

CONTENTS

1 INTRODUCTION

In this section we will present the motivations and objectives of our work. First, we will comment on the importance of medical images for diagnostics, then explore the use of image processing algorithms on these kinds of images and what problems they contend with. In the motivation subsection we will first explore the state of the art algorithms for X-ray image enhancement, then discuss the use of learning based algorithms for image enhancement outside the X-ray domain. After this we will bring examples of learning based algorithms for X-ray image enhancement and point out their limitations. To finish this subsection we explain the proposed evolutionary solution, justifying the chosen approach and the metrics used. Finally on the objectives subsection we present the overall and specific goals of the research.

Medical images are crucial for diagnosis, providing valuable information to health-care professionals for accurate and timely diagnosis, treatment planning, and monitoring of various medical conditions. Technologies such as X-rays, computed tomography (CT) scans, Magnetic resonance imaging (MRI), and ultrasound allow healthcare professionals to visualize internal body structures in detail, aiding in the identification of abnormalities, tumors, fractures, and other conditions that may not be visible through physical examination alone. These imaging technologies facilitate early detection of disease and identification of abnormalities, often before symptoms manifest. Early diagnosis can significantly improve treatment outcomes and increase the likelihood of successful interventions.

The use of image processing algorithms for classification and segmentation of medical images is explored in the literature. In (GAZDA *et al.*, 2021) the authors propose a Self-Supervised Deep Convolutional Neural Network for Chest X-Ray Classification. They used accuracy, area under the receiver operating characteristic curve (AUC), sensitivity(SEN) and specificity (SPE) to measure the success of the model. It achieved results comparable to supervised state-of-the-art algorithms. In (SHAO *et al.*, 2023) the authors present an unsupervised method for the detection of intracranial aneurysms using unsupervised dual-branch learning. They used healthy vessels detection accuracy, aneurysm vessels detection accuracy and F1 score to evaluate the model. The proposed method outperformed state-of-the-art algorithms, especially in the detection of aneurysms vessels. In (YADAV *et al.*, 2023) the authors put forward an unsupervised framework based on a Generative Adversarial Network (GAN) to classify lung diseases from chest X-ray and CT images. They used Precision, Recall, F1 Score and Accuracy to validate the model. The proposed method outperformed state-of-the-art unsupervised models

in lung disease classification. In (AHN *et al.*, 2020) the authors introduce an unsupervised CNN classifier for medical images using zero-bias convolutional auto-encoders and context-based feature augmentation. They used accuracy and area under curve (AUC) as metrics for evaluation. The proposed approach achieved better accuracy when compared to state-of-the-art unsupervised methods. In (DING *et al.*, 2022) the authors propose an unsupervised fuzzy clustering algorithm for Chest X-ray and CT images segmentation to assist COVID-19 detection. The authors used four different cluster validity indices for evaluation purposes: Davies–Bouldin, Xie-Beni, Dunn and β indexes. The proposed approach outperforms other state-of-the-art algorithms in most indexes, only being slight worst in the β index.

Medical images, like any other digital images, can be susceptible to various quality issues that impact their accuracy and utility in clinical settings. These issues can arise from equipment defects, environmental conditions, and limitations in the imaging procedure related to the patient's health condition. Problems such as noise, artifacts, blur, and contrast issues can compromise the diagnostic capabilities of both physicians and algorithms. Addressing and mitigating these image quality issues is crucial to ensure accurate diagnosis, whether conducted by physicians or algorithms. Hence, it is essential to develop new and improved image enhancement algorithms for medical images.

1.1 Motivation

Image enhancement techniques are used in the literature to improve the results of image processing algorithms. Various works have been performed to study the effect of traditional image enhancement algorithms such as Unsharp Masking (UM), High-Frequency Emphasis Filtering (HEF), and Contrast Limited Adaptive Histogram Equalization (CLAHE) on new image processing algorithms and tasks. UM is a widely used image enhancement technique to sharpen edges and increase local contrast in an image. It works by subtracting a blurred version of the original image from the original image itself, thereby enhancing edges and fine details. HEF is another image enhancement technique used to enhance the sharpness of edges in an image. It works by boosting the high-frequency components of the image, which correspond to the edges and fine details. CLAHE is an image processing technique utilized to enhance contrast in images. It extends the traditional Histogram Equalization method by adapting the process locally, which is beneficial for images with varying illumination conditions. Additionally, CLAHE incorporates a contrast limiting mechanism to prevent excessive amplification of contrast, particularly in

regions with low local contrast.

These algorithms are widely used and their utility as pre-processing tools for image processing algorithms has been explored in the literature as shown by the following works. In (WU *et al.*, 2021), the authors propose a combination of CLAHE and a dual branch network for lung nodule segmentation on computed tomography (CT) images. They used the Dice coefficients of nodule segmentation to validate their model and found that it outperforms other state-of-the-art solutions. In (RAHMAN *et al.*, 2021a) the authors explored the effect of five image enhancement algorithms on COVID-19 detection using chest X-Ray images: histogram equalization (HE), CLAHE, image complement, gamma correction, and balance contrast enhancement technique (BCET). They used accuracy, precision, sensitivity, F1-score, and specificity as metrics to evaluate the effectiveness of the proposed method. They found that the image enhancement techniques contributed to improving the COVID-19 diagnostic algorithms applied to chest X-Ray images. In (LIN *et al.*, 2020), the authors propose the use of image enhancement algorithms to enhance the accuracy of machine learning-based classifiers for chest X-ray images. They employed a two-dimensional fractional-order convolution for image enhancement and explored commonly used algorithms such as UM, HEF, and CLAHE. They evaluated the proposed method using recall, precision, accuracy and F1 score. All score indexes of the proposed classifier were higher than those of state-of-the-art solutions.

Furthermore, recent research has also introduced new image-processing algorithms for image enhancement of X-ray images. (KUMAR; BHANDARI, 2022) introduced the Automatic Tissue Attenuation-based Contrast Enhancement (ATACE) algorithm, designed to enhance X-ray images to aid in the identification of bones and organs. They used multiple image quality assessment metrics: modified measure of enhancement (MEME), tenengrad value (TEN), entropy, natural image quality evaluator (NIQE), contrast-changed image quality measure (CEIQ), and quality-aware relative contrast measure (QRCM). The proposed method outperforms state-of-the-art algorithms in MEME, TEN, QRCM, CEIQ, and entropy while being competitive in NIQE. (KUMAR *et al.*, 2021) presented the Triple Clipped Dynamic Histogram Equalization (TCDHE) algorithm, which enhanced basic features, preserved brightness, and improved contrast in medical images. This algorithm achieved good results through histogram manipulation and simple computations, relying solely on image transformations. The authors utilized Entropy as a no-reference image enhancement metric and Structure similarity index measurement (SSIM), feature similarity index measurement (FSIM), peak signal to noise ratio (PSNR), gradient magni-

tude similarity deviation (GMSD) and edge preservation index (EPI) as a full-reference image enhancement metric. The proposed method outperformed state-of-the-art algorithms on X-ray images in all metrics and on COVID-19 CT images in all metrics except entropy. These image enhancement algorithms are based on conventional image processing techniques such as spacial and frequency transformations, convolutions, and statistical functions, without making use of artificial intelligence.

The use of learning-based algorithms for image enhancement outside of the X-ray dominion is explored in the literature. A highly effective unsupervised generative adversarial network is used in (JIANG *et al.*, 2021) for low-light image enhancement. The authors used an attention-guided U-Net as the generator, a dual-discriminator to direct the global and local information, and a self-feature preserving loss to guide the training process and maintain the textures and structures. The model is compared to other competing methods using the Natural Image Quality Evaluator (NIQE) no-reference score. In (HUANG *et al.*, 2022) the authors propose a deep convolutional neural network for endoscopic image enhancement that fuses three enhanced images created by using either gamma correction, CLAHE or illumination map estimation. They created a loss function similar to Structural Similarity Index Measure (SSIM) comparing the fused image with the three starting original images. The algorithm showed better results in Entropy, Contrast Improvement Index and Average Gradient than other state-of-the-art methods.

But while there are works in the literature exploring learning-based algorithms for X-ray image enhancement they have limitations, being either limited by the use of synthetic noise or the use of simpler models when compared to Convolutional Neural Networks (CNNs). In (LUO *et al.*, 2021), the authors propose a supervised dehazing CNN for Cardiac Catheter-Based X-ray procedures to improve the quality of images obtained using low radiation dosage exams, but make use of synthetic noise to generate their dataset, which puts into doubt whether the results are representative of real-life situations. In (MALALI *et al.*, 2021) the authors propose the use of local s-Curve spacial transformations for X-Ray Mammograms contrast enhancement. The image is divided in same size patches and the gray level of each pixel is modified using a non-linear function. The function limits are determined by the maximum and minimum intensity of the local pixels. The function parameters are adjusted using a multi-objective genetic algorithm with the IQA metrics: Effective Measure of Enhancement (EME), Edge Content (EC), Feature Similarity Index Measure (FSIM) and Absolute Mean Brightness Error (AMBE) for fitness

calculation. This solution is limited by the model used, which relies on spatial transformations established by local functions.

To address this limitations, this study proposes a novel unsupervised evolutionary algorithm for image enhancement: Image Quality Assessment-based Evolutionary Network (IQAEvolNet), an evolutionary algorithm that trains a convolutional neural network for image enhancement using image quality assessment (IQA) metrics as fitness parameters. This solution would allow for the unsupervised training, without relying on synthetic noise, of convolutional neural networks (CNN) for X-ray image enhancement. Unlike the more common approach of using generative adversarial networks (GANs), an evolutionary approach is chosen. While GANs can serve this purpose, they are constrained by the discriminator's ability to accurately assess the enhanced image. As highlighted in (LEPCHA *et al.*, 2023), current state-of-the-art algorithms often lack robustness and adaptive capabilities required for real-world applications. The simpler evolutionary approach offers greater flexibility and adaptability to different applications. The proposed algorithm uses a modified single-channel version of the convolutional neural network (CNN) introduced in (ULLAH *et al.*, 2021), known as LightDehazeNet, as a base model. The top layers of the CNN have been adjusted to better suit the image enhancement task, resulting in a variant named lightdehazeTestnet.

IQAEvolNet fine-tunes the model's weights using an evolutionary strategy. The fitness of the population is determined by two no-reference image quality assessment (IQA) metrics: Entropy and Natural Image Quality Evaluator (NIQE), along with the Structural Similarity Index Measure (SSIM) between the enhanced image and the original image, which serves as a distortion metric. These IQA metrics were chosen for being commonly utilized in the literature.

The proposed solution works as follows. The base model undergoes initial training using state-of-the-art chest X-ray image enhancement algorithms as references (UM, HEF, CLAHE, ATACE, TCDHE). The resulting weights from this training serve as starting individuals for the evolutionary algorithm. An initial population is constructed by incorporating weights created at various epochs during the CNN training process. In each training round, every weight set is loaded into the model and applied to a sample from the unlabeled data set. Fitness is determined by applying no-reference image quality assessment metrics to the outcomes of each individual, along with distortion metrics assessing the changes from the original images. Individuals who demonstrate superior performance contribute to the next generation of weights

through permutation and mutation. After each epoch of training, the newly achieved solution is compared with the best previous solution. If it shows greater fitness across the entire training dataset, it supersedes the old best solution. The algorithm concludes either after reaching the absolute epoch limit or after a specified number of epochs without improvement. At the termination of the algorithm, it returns the best solution identified throughout the evolutionary process. The code is available at <https://github.com/Carlos-Alfredo/IQAEvolNet>.

1.2 Objectives

The goal of this work is to propose an unsupervised evolutionary algorithm for enhancing X-ray images. A CNN originally designed for lightweight image dehazing is employed and trained using an evolutionary approach with an unlabeled dataset. IQA metrics are utilized for fitness calculation during the evolutionary process.

The specific objectives of this work are:

- Development of an unsupervised Evolutionary Algorithm with a CNN model for Chest X-ray image enhancement: IQAEvolNet;
- Evaluate the performance of the algorithm for image enhancement tasks using no-reference IQA metrics;
- Compare the proposed algorithm with state-of-the-art algorithms.

In the next section we will discuss the theoretical background of our work, exploring the chosen algorithms and the reason for their selection. Then, on the Methodology section we detail the datasets used, explain the algorithm, and describe the experiments. Following this, on the Results section we present the enhancement results using multiple fitness equations and the classification results. Finally, on the Conclusion section we present our final considerations and the contributions of the work.

2 THEORETICAL BACKGROUND

Our work lies at the intersection of two lines of study: image enhancement and evolutionary algorithms. First, we will discuss the different approaches for image enhancement found in the literature, evaluating both learning-based and non-learning-based methods. Then, we will explore the image quality assessment algorithms used in the literature and justify the selection made for this work. Finally, we will discuss the current applications of evolutionary algorithms for various purposes such as image classification, neural network hyper-parameter optimization, and energy load forecasting.

2.1 Image enhancement without learning

For X-ray image enhancement, the most commonly used algorithms are Unsharp Masking (UM), High-Frequency Emphasis Filtering (HEF), and Contrast Limited Adaptive Histogram Equalization (CLAHE). The authors in (MUNADI *et al.*, 2020) studied the impact of image enhancement algorithms on tuberculosis detection in chest X-ray images using deep learning. They tested UM, HEF, and CLAHE and found that these algorithms improved the accuracy of classification algorithms by up to 3.37%. Similarly, in (HAYATI *et al.*, 2023), the authors studied the impact of using CLAHE as an image enhancement algorithm on diabetic retinopathy classification through deep learning. They found that while CLAHE increases the accuracy for the VGG16, InceptionV3, and EfficientNet models, it also decreases the accuracy for the ResNet34 model.

There are also more recent works proposing new image-processing algorithms for image enhancement. In (KUMAR; BHANDARI, 2022), the authors developed an Automatic Tissue Attenuation-based Contrast Enhancement (ATACE) algorithm to improve X-ray images, facilitating the identification of bones and organs. A novel Triple Clipped Dynamic Histogram Equalization (TCDHE) was proposed in (KUMAR *et al.*, 2021) to improve the basic features, brightness preservation, and contrast of medical images. By manipulating histograms and performing simple computations, the authors created an algorithm capable of enhancing X-ray and CT images. This approach achieved good results relying only on image transformations, using entropy for no-reference image enhancement quality assessment. In (BRUCE *et al.*, 2020), image enhancement techniques are used to improve ultrasound images of blood flow. In (LIU; TIAN, 2020), the authors used an algorithm based on fractional differential to enhance bone

X-ray images while preserving detailed information on the texture.

From the options of image-processing algorithms for image enhancement our work will use UM, HEF, CLAHE, ATACE and TCDHE. We will use they as starting points for our algorithm. Their algorithms are described in subsections 2.1.1, 2.1.2, 2.1.3, 2.1.4, and 2.1.5.

2.1.1 Unsharp Masking

UM is a widely used image enhancement technique to sharpen edges and increase local contrast in an image. It works by subtracting a blurred version of the original image from the original image itself, thereby enhancing edges and fine details.

The algorithm cam be divided in 4 steps. First, a Gaussian filter is applied to create a blurred version of the original image. Equation 2.1 express this step, where $G(x,y)$ is the blurred image, $I(x,y)$ is the original image, $*$ is a convolution operation, and $H(x,y)$ is a Gaussian blur kernel. Then, the blurred image is subtracted from the original image to obtain the high-pass filtered image, which emphasizes edges and fine details. Equation 2.2 shows this step, where $E(x,y)$ is the high-pass image. Next, the high-frequency components are amplified by scaling the high-pass image using a factor k, as shown in Equation 2.3. Then, the amplified high-frequency component is added back to the original image, as shown in Equation 2.4, where $S(x,y)$ is the sharpened image.

$$G(x,y) = I(x,y) * H(x,y) \tag{2.1}$$

$$E(x,y) = I(x,y) - G(x,y) \tag{2.2}$$

$$E'(x,y) = k \times E(x,y) \tag{2.3}$$

$$S(x,y) = I(x,y) + E'(x,y) \tag{2.4}$$

2.1.2 High-Frequency Emphasis Filtering

HEF is another image enhancement technique used to enhance the sharpness of edges in an image. It works by boosting the high-frequency components of the image, which correspond

to the edges and fine details. The first step of HEF is to convert the image to the frequency domain, as shown in Equation 2.5, where $F(u,v)$ is the frequency domain representation of the image, FFT is the Fourier transform, and (u,v) are the frequency domain coordinates. The second step is to boost the high-frequency components of the image using a high-pass filter. This is shown in Equation 2.6, where $F'(u,v)$ is the modified frequency domain representation of the image with boosted high frequencies, and $H(u,v)$ is the high-pass filter. The high pass filter is calculated using Equations 2.7 and 2.8, where D_0, K_1 and K_2 are manually adjustable parameters. The last step is to convert the image back to the spatial domain using the inverse Fourier transform. Equation 2.9 shows this process, where $I'(x,y)$ is the enhanced image in the spacial domain, and $IFFT$ is the inverse Fourier transform.

$$F(u,v) = FFTI(x,y) \tag{2.5}$$

$$F'(u,v) = H_{HFE}(u,v) \times F(u,v) \tag{2.6}$$

$$H(u,v) = 1 - e^{-\frac{(\frac{u-M}{2})^2+(\frac{v-N}{2})^2}{2\times D_0^2}} \tag{2.7}$$

$$H_{HFE}(u,v) = K_1 + K_2 H(u,v) \tag{2.8}$$

$$I'(x,y) = IFFTF'(u,v) \tag{2.9}$$

2.1.3 Contrast Limited Adaptive Histogram Equalization

CLAHE (Contrast Limited Adaptive Histogram Equalization) is an image processing technique used to enhance the contrast in images. It is an extension of the traditional Histogram Equalization method. While Histogram Equalization adjusts the intensity levels of an entire image based on its overall histogram, CLAHE adapts the process locally, making it particularly useful for images with varying illumination conditions. Additionally, it includes a contrast

limiting mechanism to prevent over-amplification of contrast, especially in regions with low local contrast.

The histogram equalization algorithm can be divided in 5 steps. First, the Histogram on the image, $h(r_k)$, is computed as shown in Equation 2.10, where $h(r_k)$ is the histogram value for the intensity level r_k and n_k is the number of pixels with intensity r_k.

$$h(r_k) = n_k \tag{2.10}$$

Then the cumulative distribution function (CDF), which represents the cumulative sum of histogram values, is computed as shown in Equation 2.11, where $cdf(r_k)$ is the cumulative distribution function at intensity level r_k.

$$cdf(r_k) = \sum_{i=0}^{k} h(r_i) \tag{2.11}$$

After that, the CDF is normalized to the range $[0,1]$ by diving each value by the total number of pixels in the image as shown in Equation 2.12, where N is the total number of pixels in the image.

$$cdf\_normalized(r_k) = \frac{cdf(r_k)}{N} \tag{2.12}$$

In the next step, the intensity transformation function, $T(r_k)$, is created by scaling the normalized CDF to the maximum intensity value (L-1) as shown in Equation 2.13, where L is the total number of intensity levels.

$$T(r_k) = \lfloor (L-1) \times cdf\_normalized(r_k) \rfloor \tag{2.13}$$

Finally, the intensity transformation function is applied to each pixel in the original image as shown in Equation 2.14, where r is the original intensity value, s is the transformed intensity value and $T(r)$ is the intensity transformation function.

$$T(r_k) = \lfloor (L-1) \times cdf\_normalized(r_k) \rfloor \tag{2.14}$$

The CLAHE algorithm has 5 steps. First, the image is divided into non-overlapping patches, with the patch size being one of the parameters set by the user. Then, on each path,

histogram equalization is applied. After that it is introduced a contrast limiting mechanism to avoid over-amplification. This is typically done by clipping the scaling factor based on a specified limit L.

$$clip(x,a,b) = \begin{cases} a & \text{if } x < a \\ x & \text{if } a \leq x \geq b \\ b & \text{if } x > b \end{cases} \tag{2.15}$$

The scaling factor k_i for the i-th tile is computed as:

$$k_i = clip\left(\frac{L}{max(cdf\_normalized_i) - min(cdf\_normalized_i)}, 0, 1\right) \tag{2.16}$$

In the next step, the transformed intensity values are adjusted using the computed scaling factor k_i as shown in the Equation 2.17, where s'_i is the final adjusted intensity value.

$$s'_i = \lfloor k_i \times s_i \rfloor \tag{2.17}$$

Finally, the adjusted tiles are combined to obtain the final enhanced image.

2.1.4 Automatic tissue attenuation-based contrast enhancement

The ATACE algorithm has been proposed on (KUMAR; BHANDARI, 2022) to improve the contrast of X-ray images. The first step of the algorithm is to normalize the input image. The goal of the algorithm is to find the detail and removable components. Equation 2.18 shows this step, where $I(nor)$ is the normalized image, $I(y)$ is the input image, $I_{max}(y)$ is the maximum value of the input image, $D(y)$ is the detail component, and $R(y)$ is the removable component. After this, maps of local maximum and local minimum are created by using the local region around the pixel y, as shown in Equations 2.19 and 2.20, where $G(y)$ is the local maximum component, $T(y)$ is the local minimum component, and Lo_y is the local region around the pixel y. To calculate the removable content we multiply the local minimum by a removal factor. Equations 2.21 and 2.22 illustrate this, where β is the removal factor, exp is the exponential function, and $var()$ is the variance. Another parameter, $\psi(y)$, is used to adjust the image for brightness consistency. It is calculated using Equation 2.23. Finally, the enhanced version is calculated using Equation 2.24.

$$I_{nor}(y) = \frac{I(y)}{I_{max}} = D(y) + R(y) \tag{2.18}$$

$$G(y) = max_{x \in Lo_y} I_{nor}(y) \tag{2.19}$$

$$T(y) = min_{x \in Lo_y} I_{nor}(y) \tag{2.20}$$

$$R(y) = \beta \times T(y) \tag{2.21}$$

$$\beta(y) = exp\lfloor -\frac{G(y) \times var(T(y))}{T(y)} \rfloor \tag{2.22}$$

$$\psi(y) = \frac{log\left[1 - R(y) \times \left(\frac{1}{G(y)} - 1\right)\right]}{log[G(y)]} \tag{2.23}$$

$$E(y) = \frac{I_{nor}(y) - R(y)}{G(y)^{\psi(y)} - R(y)} \tag{2.24}$$

2.1.5 Triple clipped dynamic histogram equalization

TCDHE was proposed in (KUMAR *et al.*, 2021) to improve the basic features, brightness preservation and contrast of medical images. The first step of the algorithm is to calculate the mean and standard deviation of the input image using Equations 2.25 and 2.26. Using this information we will divide the histogram of the input image into three equal partitions. The partitioning points s_1 and s_2 are calculated using Equations 2.27 and 2.28. Each of the sub-histogram is clipped using a threshold value calculated from the mean of the total pixels of each sub-histogram. This is expressed in Equation 2.29. The clipped histogram are calculated using Equation 2.30, where $h(k)$ is the histogram of the input image, and $h_c(k)$ is the clipped histogram. The PDF and CDF of every sub-histogram are calculated using Equations 2.31 and 2.32, where N_p is the whole number of pixels in every clipped sub-histogram. Then every partition is mapped into a new dynamic range to balance the process of equalization using Equations 2.33 to 2.36, where $L-1$ is the maximum intensity value of the n-bit image. After

this, each transfer function is calculated using the dynamic ranges and the CDF using Equations 2.37 to 2.39. The enhanced image, $I'(m,n)$, is calculated using the Equation 2.40.

$$\mu = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} X(u,v) \tag{2.25}$$

$$\sigma = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} [X(u,v) - \mu]^2 \tag{2.26}$$

$$s_1 = \mu - 0.43\sigma \tag{2.27}$$

$$s_2 = \mu + 0.43\sigma \tag{2.28}$$

$$T = \frac{1}{I_{max}} \times \sum_{k=0}^{I_{max}} h(k) \tag{2.29}$$

$$h_c(k) = \begin{cases} h(k) & h(k) < T \\ T & h(k) \geq T \end{cases} \tag{2.30}$$

$$PDF(k) = \begin{cases} \frac{h_c(k)}{N_p} & 0 \geq k < I_{max} \end{cases} \tag{2.31}$$

$$CDF(k) = \begin{cases} \sum_{k=0}^{I_{max}} PDF(k) & 0 \geq k < I_{max} \end{cases} \tag{2.32}$$

$$p_0 = 0 \tag{2.33}$$

$$p_1 = \frac{s_1 - s_0}{s_3 - s_0 + 1} \times (L-1) \tag{2.34}$$

$$p_2 = \frac{s_2 - s_1}{s_3 - s_0 + 1} \times (L-1) + p_1 \tag{2.35}$$

$$p_3 = \frac{s_3 - s_2}{s_3 - s_0 + 1} \times (L-1) + p_2 = L-1 \tag{2.36}$$

$$TF_{lower} = (p_1 - 1) \times CDF_{Lower} \tag{2.37}$$

$$TF_{Middle} = p_1 + (p_2 - 1 - p_1) \times CDF_{Middle} \tag{2.38}$$

$$TF_{upper} = p_2 + (p_3 - p_2) \times CDF_{Upper} \tag{2.39}$$

$$I'(m,n) = \begin{cases} TF_{lower}(I(m,n)) & I(m,n) < s_1 \\ TF_{middle}(I(m,n) - s_1) & s_1 \geq I(m,n) < s_2 \\ TF_{upper}(I(m,n) - s_2) & s_2 \geq I(m,n) \end{cases} \tag{2.40}$$

The original and enhanced image are then decomposed using Discrete Wavelet Transform (DWT). DWT is commonly used in image processing for tasks like compression, denoising, and feature extraction. The DWT operates by decomposing the image into approximation and detail coefficients across different scales and orientations. It involves a series of filtering and downsampling operations. Equation 2.41 and 2.42 shows the decomposition, where LL and LL' are the approximation coefficients of the original and enhanced image, and LH, HL, HH and LH', HL', HH' represent the detailed coefficients of the original and enhanced image.

$$LL, LH, HL, HH = DWT(I(m,n)) \tag{2.41}$$

$$LL', LH', HL', HH' = DWT(I'(m,n)) \tag{2.42}$$

After this decomposition, the lower frequency contents LL and LL' are decomposed using singular value decomposition (SVD), expressed in Equation 2.43, where B is the image to be decomposed, U_B and V_B are orthogonal matrices, T is the transpose operator, and Σ is identified as the matrix of a singular value, which has the intensity value of the input image. LL and LL' are decomposed to $U_{LL}, \Sigma_{LL}, V_{LL}$ and $U_{LL'}, \Sigma_{LL'}, V_{LL'}$, respectively. A improvement factor, ξ, is calculated using Equation 2.44. The normalized matrix of a singular value is calculated using Equation 2.45, where β is an empirically adjustable parameter between 0.05 and 0.95. The normalized approximation coefficient LL_N is calculated using Equation 2.46.

$$B = U_B \Sigma_B V_B^T \tag{2.43}$$

$$\xi = \frac{max(U_{LL'}) + max(V_{LL'})}{max(U_{LL}) + max(V_{LL})} \tag{2.44}$$

$$\Sigma_N = (\beta \xi \Sigma_{LL}) + \left((1-\beta) \frac{1}{\xi} \Sigma_{LL'} \right) \tag{2.45}$$

$$LL_N = U_{LL'} \Sigma_N V_{LL'}^T \tag{2.46}$$

The detailed coefficients are normalized using spatial frequency (SF). SF is the amount of frequency components of the image and it is calculated using Equations 2.47 to 2.49, where f_r, f_c, f_s are row, column, and spatial frequencies for a $M \times N$ image respectively.

$$f_r = \sqrt{\frac{1}{MN} \sum_{M-1}^{u=0} \sum_{N-1}^{v=1} [I_f(u,v) - I_f(u,v-1)]^2} \tag{2.47}$$

$$f_r = \sqrt{\frac{1}{MN} \sum_{V-1}^{v=0} \sum_{M-1}^{u=1} [I_f(u,v) - I_f(u-1,v)]^2} \tag{2.48}$$

$$\sqrt{f_r^2 + f_c^2} \tag{2.49}$$

The normalized SF of LH and LH' is calculated using Equations 2.50 and 2.51, where $S_N(LH)$ and $S_{N'}(LH')$ are the normalized detailed coefficient of the input and enhanced image respectively. Both are them fused using Equation 2.52 to obtain a new LH sub-band image LH_N.

$$S_N(LH) = \frac{S_{LH}}{S_{LH} + S_{LH'}} \tag{2.50}$$

$$S_N(LH') = \frac{S_{LH'}}{S_{LH} + S_{LH'}} \tag{2.51}$$

$$LH_N = S_N(LH) \times LH + S_N(LH') \times LH' \tag{2.52}$$

The other sub-band images are calculated by fusing their respective normalized detailed coefficients. Finally, the inverse DWT is applied over the new sub-band images as shown in Equation 2.53.

$$I_{ENHANCED} = IDWT(LL_N, LH_N, HL_N, HH_N) \tag{2.53}$$

2.2 Image enhancement with supervised learning

Convolutional neural networks (CNNs) are powerful tools for image processing, enabling the leverage of large datasets and learning approaches to solve complex problems. In this subsection, we will discuss the operation of CNNs and the state-of-the-art supervised image enhancement algorithms.

2.2.1 Convolutional neural networks

CNNs are a class of deep learning models specifically designed for processing structured grid data, such as images. They leverage convolutional layers to automatically and adaptively learn hierarchical features from the input data. The core operation in a CNN is the convolution operation, which involves sliding a filter (also known as a kernel) over the input data to perform element-wise multiplication and summation. The mathematical representation of

the 2D convolution operation is shown on Equation 2.54, where I is the input image, K is the convolution kernel, (i, j) are the pixel coordinates in the output feature map and (m,n) are the pixel coordinates in the filter.

$$(I * K)(i, j) = \sum_m \sum_n I(i+m, j+n) \times K(m,n) \tag{2.54}$$

The basic building blocks of a CNN are convolutional layers, activation functions, pooling operations, and fully connected layers. A convolutional layer typically involves multiple filters applied to the input to generate multiple channels in the output feature map. The output of a convolutional layer is obtained by applying the convolution operation and adding a bias term. Equation 2.55 shows how this calculation is done, where O is the output feature map, σ is the activation function, (i, j, k) are the pixel coordinates in the filter, (m,n) are the pixel coordinates in the filter, l represents the input channels and $K(m,n,l,k)$ is the weight of the filter.

$$O(i, j, k) = \sigma(\sum_m \sum_n \sum_l I(i+m, j+n, l) \times K(m,n,l,k) + b_k) \tag{2.55}$$

Activation functions play a crucial role in Convolutional Neural Networks (CNNs) by introducing non-linearity into the network. This non-linearity is essential for the network to learn and approximate complex relationships within the data. Common activation functions used in CNNs include Rectified Linear Unit (ReLU), sigmoid, and hyperbolic tangent (tanh). The ReLU activation function is expressed on Equation 2.56.

$$ReLU(x) = max(0, x) \tag{2.56}$$

Pooling layers are often used to down-sample the spatial dimensions of the feature map. They contribute to the efficiency, robustness, and generalization capability of CNNs by reducing spatial dimensions, introducing translation invariance, and enhancing the network's ability to capture hierarchical features. The max pooling operation is expressed on Equation 2.57, where P is the output of the pooling layer, and (i, j, k) are the pixel coordinates and channel index in the output.

$$P(i, j, k) = max_{m,n} O(2i+m, 2j+n, k) \tag{2.57}$$

After several convolutional and pooling layers, a CNN often ends with one or more fully connected layers. The output of a fully connected layer is obtained by applying a linear transformation followed by an activation function. Equation 2.58 shows how it is calculated, where x is

the input vector, W is the weight matrix, b is the bias vector, and σ is the activation function.

$$FC(x) = \sigma(Wx + b) \tag{2.58}$$

The network learns the parameters (weights and biases) during training using optimization algorithms such as stochastic gradient descent (SGD) to minimize a specified loss function. The learned features enable the network to perform tasks such as image classification, object detection, and segmentation.

2.2.2 State of the art supervised image enhancement

The popularization of convolutional neural networks (CNNs) for image processing tasks has led to an explosion of new approaches to addressing image enhancement challenges. New models have been developed to handle different contexts and requirements by leveraging paired datasets. In the survey (JAMIL *et al.*, 2023), various examples of image denoising, a subset of image enhancement, using vision transformers are discussed. Each of these examples was designed for different types of images, such as computed tomography (CT) scans, natural images, and hyperspectral images. In another work (YANG *et al.*, 2022a), authors propose a supervised image enhancement model for computational tomography of asphalt to enable a better understanding of its internal structure.

Low-light image enhancement is one of the main areas explored in the literature. Authors in (FAN *et al.*, 2022) create an end-to-end low-light image enhancement model capable of mining the deep multi-scale features in the image. The authors use the SE-Res2block, based on Res2net, to give the model powerful feature representation capabilities; the U-Net as the backbone of the model; and a combination of Structural Similarity Index Measure (SSIM) and perceptual loss as the loss function. A progressive dual network low exposure image enhancement model is developed in (ZHOU *et al.*, 2022) based on Retinex theory. The network takes the low-exposure image as the input, uses convolution kernels of different scales for feature extraction, and finally learns the illumination map in the Retinex model. Then, the illumination map is substituted into the Retinex model to calculate the brightness-enhanced image. Then, aiming at the noise amplification problem in the enhancement process, the enhanced image is passed through an image-denoising network to obtain the final enhancement result.

In the literature, low-light image enhancement is a prominent area of exploration. In the study (FAN *et al.*, 2022), researchers developed an end-to-end low-light image enhancement

model capable of extracting deep multi-scale features from images. The model incorporates the SE-Res2block, based on Res2net, to provide powerful feature representation capabilities. Additionally, it utilizes the U-Net as the backbone and employs a combination of Structural Similarity Index Measure (SSIM) and perceptual loss as the loss function. Another study (ZHOU *et al.*, 2022) presents a progressive dual-network approach for low-exposure image enhancement, based on the Retinex theory. The network takes the low-exposure image as input, extracts features using convolution kernels of different scales, and learns the illumination map in the Retinex model. Subsequently, the illumination map is used to calculate the brightness-enhanced image. To address the noise amplification issue during enhancement, the enhanced image undergoes denoising through an image-denoising network, resulting in the final enhancement outcome.

Another area extensively explored in the literature is underwater image enhancement. A CNN-based underwater image enhancement model (WANG *et al.*, 2021) uses both the RGB and HSV color spaces. The RGB block implements fundamental operations. The HSV block implements global adjustments. The attention map block combines the RGB and HSV block output images. An underwater image enhancement framework based on transfer learning (ZHANG *et al.*, 2023b), uses a domain transformation module and an image enhancement module to perform color correction and image enhancement, respectively.

Beyond that, the literature already explores image enhancement models for some types of medical images. The survey (SHAMSHAD *et al.*, 2023) points out the impressive success achieved by using vision transformers in enhancing Low-Dose Computed Tomography. In (YE *et al.*, 2021) the authors combine both supervised and unsupervised learning for a low-dose CT Image Reconstruction framework. In (QI *et al.*, 2021) the authors propose a a deep neural network to enhance Plane Wave Imaging, an ultra-fast medical ultrasound imaging mode whose image quality is severely degraded in comparison to standard ultrasound, while maintaining a high frame rate. The authors in (WU *et al.*, 2024) proposed a supervised model for transforming low resolution medical images in high resolution ones is to improve disease diagnostic.The authors propose a super-resolution algorithm which is adaptable to the type of medical image. They do this by using a improved version of the Adam optimizer where each parameter has the ability to measure it's importance for the task at hand. In (CHOKCHAITHANAKUL *et al.*, 2022) the authors propose a augmentation method for Tuberculosis Chest X-ray datasets by using a contrast enhancement technique that automatically identifies the lung region and normalizes the image accordingly. They use this enhancement technique to generate images of

different levels of quality and, together with other augmentation techniques such as rotation and brightness, improve the robustness of the data. Authors create a chest X-ray contrast enhancement model capable of adjusting the desired local level of contrast. The authors mix CLAHE with a CNN-based residual learning network to adjust the level of contrast locally. They validate the results through both image enhancement metrics and classification algorithms performance enhancement.

2.3 Image enhancement with unsupervised learning

The main problem with supervised approaches is the construction of the labeled datasets. For image enhancement tasks there is the problem with its subjective nature. Different individuals may have different opinions on what constitutes a well-enhanced image. Beyond that, obtaining large-scale datasets with paired images can be resource-intensive, particularly in cases where manual enhancement is involved. There is the option of acquiring high-quality images that represent the same ground truth of low-quality images but, for medical images, that would require a pair of low-quality/high-quality images from the same patient acquired in a small time interval. In the survey (LEPCHA *et al.*, 2023) the authors points out that numerous image enhancement algorithms rely on the generation of synthetic data for their training. They explain that this tactic can produce good results in synthetic data but doesn't translate in good results in real data and, therefore, it is of vital importance the development of unsupervised solutions for image enhancement applications. For these reasons there has been extensive research done in the literature on unsupervised learning models.

There is an extensive literature in low-light and underwater image enhancement using unsupervised learning algorithms. In (HU *et al.*, 2021) the authors propose an unsupervised learning-based low-light image enhancement for indoors images. They divide the problem in two stages. In the first stage they use tone mapping for illumination enhancement. While this stage is simple and can improve the illumination of the image, it also amplifies noise. The second stage is used to minimize this effect. In this stage the model trains an encoder-decoder network using a comprehensive loss function with a combination of loss of image content, perceptual quality, total variation, and adversarial loss. In (KANDULA *et al.*, 2023) the authors propose an unsupervised low-light enhancement network using context-guided illumination-adaptive norm. First the network multiplies each pixel of the input image by an individual amplification factor. Then the networks uses a decoder-encoder structure to create an improved enhanced

image with a data distribution more similar to well-lit images. In (LI *et al.*, 2023) the authors propose a unsupervised method for mixed-exposure low-light enhancement called LEES-Net. They combined a spacial transformation curve for pixel intensity with an attention mechanism to enhance the image.

In (LUO *et al.*, 2024) the authors propose a pseudo-supervised low-light image enhancement with mutual learning. They first use a quadratic curve to create a reference enhanced image from a low-light image. Both images are then fed to identical networks. Each branch learns from the other which allows them to produce the final result. In (JIANG *et al.*, 2022) the authors propose an Unsupervised Decomposition and Correction Network for low-light image enhancement. They train their model using the input image and it's histogram equalization (HE) counterpart. The model is then optimized using hybrid no-reference quality-aware losses and inter-consistency constraints between the original image and the HE enhanced pair. In (LIN *et al.*, 2024) the authors propose an unsupervised low-light image enhancement algorithm using paired CNNs with distinct activation functions. They make use of 5 loss functions, each with a different goal. For quality loss they introduce a novel Bezier loss function that measures the mean difference between the intensity of the predicted image and their intensity if the image was transformed using a Bezier curve.

In (GUO *et al.*, 2022) the authors propose an unsupervised learning-based low-light image enhancement method for catenary, the electrical cable that is used to transmit electrical energy to electric trains, images. This would improve the performance of non-contact catenary detection systems, making them less susceptible to errors caused by environment factors and equipment defects. The unsupervised method is a Zero-Reference Learning with Exposure Control, Spatial Consistency and Illumination Smoothness losses. In (DING *et al.*, 2024) the authors propose a Unsupervised Unified Image Dehazing and Denoising Network for outdoor images. In (YAN *et al.*, 2023) the authors point out that the lack of paired underwater and in-air image datasets difficulties the use of supervised methods and propose an unsupervised architecture for underwater image restoration.

As far as medical images, there are works proposing unsupervised algorithms for low-dose computed tomography (CT) enhancement. In (LI *et al.*, 2022) the authors propose a low-dose CT reconstruction algorithm with an unsupervised training process based on a Gaussian mixture model to characterize the noise distribution in CT images. Still in medical applications, in (XIE *et al.*, 2023) the authors propose an unsupervised multi-modal medical image fusion

algorithm. Multi-modal medical image fusion is a tool for combining the information of multiple images into a more clearer and meaningful image. This helps in the diagnostic process. The authors use transformers to capture global dependencies and CNN to capture local dependencies.

One of the main tools used in unsupervised image enhancement tasks are generative adversarial networks (GANs). GANS consist of two neural networks, a generator and a discriminator, which are trained simultaneously through adversarial training. The generator creates realistic data, and the discriminator evaluates the authenticity of that data. The GAN framework involves training the generator and discriminator to optimize their respective objectives. The training process involves a min-max game, where the generator tries to generate realistic samples to fool the discriminator, while the discriminator aims to correctly distinguish between real and generated samples. The training process advances iteratively, with the generator and discriminator updating their parameters in an adversarial manner. The generator learns to produce realistic samples, while the discriminator becomes better at distinguishing between real and generated samples. The equilibrium point is reached when the generator produces samples that are indistinguishable from real samples, and the discriminator cannot reliably differentiate between the two.

In (SUN *et al.*, 2020) the authors created a generative adversarial network to to generate a normal-looking medical image from an abnormal-looking one with the goal of assisting lesion segmentation or classification without having to rely on labeled datasets. The authors focus on magnetic resonance imaging and computational tomography. An enhancement method based on GAN developed in (XU *et al.*, 2022), to effectively enhance low-quality underwater images. The authors use both a global and a local discriminator to decide if the generated image is true or false. The discriminator uses the Patch GAN method. The loss function is the WGAN-GP and the DE-GAN is used as the counter loss. The authors create a GAN that accomplishes unpaired Low-light image enhancement through Self-Similarity Contrastive Learning (SSCL). The SSCL allows for the differentiation between the low-light domain and the normal-light domain. The dual illumination perception module captures spatial and global information and helps the network achieve a more realistic result. This method uses unsupervised learning but was only tested in a very specific context of aerial images and requires confirmation if it can be used in other contexts. In (WANG *et al.*, 2023) the authors propose a multi-modal medical image fusion algorithm based on CycleGAN with weakly supervised training. The algorithm generates CT images from magnetic resonance (MR) images. It also creates enhanced MR images that are

fused with the synthetic CT images and the positron emission tomography images to accomplish the tri-model fusion.

2.4 Image quality assessment

Image quality assessment (IQA) metrics are an important part of the validation of image enhancement algorithms. They are algorithms that quantify the quality of enhanced images. These algorithms can be divided into two groups: reference and no-reference IQA metrics. Reference IQA metrics are used when you have access to low-quality/high-quality paired dataset. In those cases, the algorithms basically measure the error between the enhanced image and the high-quality reference image. The most commonly used of this algorithms are the Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM). No-reference IQA metrics are used when you don't have access to reference images. In those cases, the algorithms try to use the information contained in the enhanced image to quantify their quality. These methods can be sub divided in image processing approaches and statistical approaches. In the image processing approaches the algorithm quantification relies solely on the image. Examples of this approach are Entropy, Average Gradient, and Contrast Improvement Index. In the statistical approaches the algorithms uses the statistical information of a big sample of images to calculate the "distance" between the enhanced image and the typical image. They assume that images with statistical information more different to the average are of worse quality. Examples of this approach are Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE), Perceptual Image Quality Evaluator (PIQE), and Natural Image Quality Evaluator (NIQE).

These IQA metrics are widely used in the literature. In (AURANGZEB *et al.*, 2021) the authors propose a modified particle swarm optimization for CLAHE parameter tuning. Their goal is to improve the sensitivity, specificity and accuracy of supervised semantic segmentation algorithms for retinal images. They used an objective function for parameter evaluation based on the number of edges obtained by the Sobel operator, the entropy of the enhanced image and the sum of the pixels of the edge image created by the Sobel operator. They used Entropy and SSIM for quantitative evaluation. In (KUMAR; BHANDARI, 2023) the authors propose a method for unsupervised classification of images based on their under/well/over-exposed illumination regions. They first classify the super-pixels, patches, on their contrast and intensity using pure image processing techniques. For measuring the contrast they use entropy. In (ZHU *et al.*, 2021) the authors propose a dehazing algorithm based on artificial multi-exposure image fusion

for outdoor images affected by poor weather conditions. They first produce multiple gamma-corrected images using different coefficients. These images are then decomposed into global and local components. These components are used to calculate pixel-wise weight maps. These weight maps are used to generate a dehazed fused image from the gamma-corrected images. The dehazed image is then balanced for image luminance and color saturation. They used Entropy for no-reference image enhancement quantified analysis.

In (YU *et al.*, 2023) the authors propose a fuzzy self-guided GAN for medical image enhancement. They use structure preservation and illumination distribution to calculate the loss function. They point out that both metrics act in opposition to each other, with the illumination distribution favoring illumination uniformity and the structure preservation favoring information preservation, and that opposition improves the final result by not producing a model that leans too much on neither tendency. They used entropy, BRISQUE, NIQE and average gradient for quantitative analysis. The experiment was done in a con-focal microscopic image dataset. In (HUANG *et al.*, 2022) the authors propose a deep unsupervised endoscopic image enhancement algorithm based on multi-image fusion. They propose a deep convolutional neural network that fuses three enhanced images created by using either gamma correction, CLAHE or illumination map estimation. They created a loss function similar to SSIM comparing the fused image with the three starting original images. The algorithm showed better results in Entropy, Contrast Improvement Index and Average Gradient than other state-of-the-art methods. In (SHI *et al.*, 2022) the authors propose an unsupervised algorithm for low-light image enhancement. The loss is calculated using a combination of structural similarity, color consistency and naturalness discriminator modules. They used NIQE for no-reference quantitative effectiveness analysis.

A highly effective unsupervised generative adversarial network (JIANG *et al.*, 2021) is used for low-light image enhancement. The authors used an attention-guided U-Net as the generator, a dual-discriminator to direct the global and local information, and a self-feature preserving loss to guide the training process and maintain the textures and structures. The model is compared to other competing methods using the NIQE no-reference score. A novel and versatile bi-directional GAN (StillGAN) proposed on (MA *et al.*, 2021) for medical image quality enhancement is capable of imposing constraints on structure and illumination. StillGAN treats low- and high-quality images as two distinct domains and introduces local structure and illumination constraints for learning both overall characteristics and local details. The method has an unsupervised training process and has achieved good results but was only tested

in very specific test scenarios (Corneal Microscopy, Color Fundus, Endoscopy) and was not stressed in more noisy environments such as ultrasound and X-ray. It uses NIQE, BRISQUE and PIQE for no-reference quality assessment. In (PRAJAPATI *et al.*, 2021) the authors propose a unsupervised super-resolution algorithm using generative adversarial network. The supervised methods usually make use of pairs of high resolution images with low resolution images created using known algorithm such as bi-cubic down-sampling. This approach does not take into account the multiple degradation factors presented in real-world data. They bench-marked the performance of their method using both reference-based scores (Perceptual Index, Root Mean square Error, and Learned Patch Image Perceptual Score) and no-reference-based scores (NIQE, BRISQUE and PIQE).

As the goal of our work is to propose an unsupervised image enhancement algorithm for Chest X-ray images we will not have ground truth images available. Therefore, the image enhancement results will be validated using no-reference IQA metrics. We chose to use Entropy and NIQE in our experiments because they are the most present in the literature for each of the no-reference IQA metrics subgroups. Beyond that, we will also make use of PSNR and SSIM as measures of distortion between the original image and the enhanced image.

2.5 Evolutionary algorithms

Evolutionary algorithms are a family of optimization algorithms inspired by the process of natural selection. These algorithms are used for solving optimization and search problems, where the goal is to find the optimal solution from a set of possible solutions. The basic idea is to mimic the process of biological evolution, where individuals in a population undergo selection, crossover (recombination), and mutation over multiple generations.

In evolutionary algorithms, potential solutions to the optimization problem are represented as individuals or candidate solutions. A fitness function is used to evaluate how well an individual performs the task or solves the optimization problem. The goal is usually to maximize or minimize the fitness function. A population consists of a collection of individuals. Each iteration of the algorithm, also referred as generation, the population undergoes a selection process. Selection involves choosing individuals from the current population to serve as parents for the next generation. The probability of selecting an individual can be based on its fitness value. In most cases, the individuals are sorted by their fitness and the ones with the higher in the list are selected. After selection, the selected individuals undergo a crossover or recombination

process. The crossover process combines genetic material from two parents to create one or more offspring. The individuals created by this process are included in the population. Next, the new population is subjected to mutation. Mutation introduces random changes to individuals to explore new regions of the solution space. The algorithm terminates after a certain number of generations or when a termination condition is met.

In the literature, evolutionary algorithms are used to deal with a wide range of problems. One of the main applications is for CNN hyper-parameter optimization. In (ZHANG *et al.*, 2021) the authors proposed an Evolutionary Algorithm for Hyper-parameter optimization of multilevel convolutional neural network for lung nodule classification. They used a evolutionary strategy to efficiently search for hyper-parameter configurations that improve the classification results. The results achieved were better than other approaches such as random search and grid search. In (LU *et al.*, 2021) the authors propose an evolutionary algorithm for automated design of classification convolutional neural networks (CNNs). They point out that most automated design solutions focus only on optimizing classification performance or only for one deployment scenario. Most solutions also require elevated amounts of computational resources. Their proposed evolutionary algorithm addresses this issues by progressively recombining and modifying the hyper-parameters of the CNN based on the fitness analysis of multiple target objectives. In (WEN *et al.*, 2021) the authors propose a two-stage evolutionary neural architecture search for transfer learning (EvoNAS-TL). First, they search for a appropriate architecture. Then, they fine-tune the structure with network pruning. In (LIN *et al.*, 2022) the authors propose an evolutionary architectural search method for GAN architecture design which simultaneously train the weights from the model. In (ZHOU *et al.*, 2021) the authors propose an evolutionary algorithm for deep neural networks compression. They model filter pruning as a multi-objective optimization problem, with the number of filters and error rate as conflicting results. They use an evolutionary approach to find the optimal trade-off point between both objectives.

Another application is as an alternative to standard gradient training of supervised models. In (YANG *et al.*, 2022b) the authors propose a gradient-guided evolutionary training method for deep neural networks. They point out that gradient-based approaches have problems with local optima and saddle points as they lack exploration ability. In this context, evolutionary solutions represent an alternative with higher exploration capacity and insensitivity to local optima. In (JALALI *et al.*, 2021) the authors created an evolutionary-based deep convolutional neural network model for energy load forecasting. In (TIAN *et al.*, 2021) the authors propose

an multi-populational evolutionary algorithm for finding multiple optimal solutions of large-scale multi-modal multi-objective problems. They use their method to train, in a supervised manner, a CNN for image classification. They found that the ensemble model of the multiple solutions performs better than other state of the art methods. In (ZHOU *et al.*, 2023) the authors propose a supervised evolutionary algorithm to train a Dynamic Multi-objective Recommendation framework for Sequential Recommendation. They point out the previous methods focus only on recommendation accuracy which can cause homogenization and redundancy. To improve this they propose a Multi-objective optimization approach where not only the accuracy is taken into account in the model training. To solve this multi-objective optimization they first train a model using only short-term recommendation accuracy and then introduce a evolutionary algorithm with supervised learning to enhance the model.

Evolutionary algorithms are also used for unsupervised multi-objective optimization tasks. In (HADIKHANI *et al.*, 2024) the authors propose an algorithm for unsupervised human activity discovery. They point out that most works in Human Activity Discovery are done with supervised methods but producing labeled data is expensive. To solve this, they created a multi-objective particle swarm optimization clustering algorithm with Gaussian mutation and game theory.

To the best of our knowledge this work is the first to explore the use of Evolutionary algorithms for unsupervised CNN training for image enhancement tasks.

3 METHODOLOGY

In this section we will present the IQAEvolNet algorithm, the datasets and the experimental procedures. First we provide an overview of our algorithm. Then on the Statistical metrics evaluation subsection we present the algorithm used for the metrics employed on the algorithm. Following this, on the LightdehazeTestnet subsection we explain the Convolutional neural network used on our algorithm. In sequence, on the IQAEvolNet subsection we describe the algorithm itself. Then, on the Datasets subsection we described all datasets used on our experiments. Finally, on the Classification test subsection we describe how the classification experiment was done.

The goal of IQAEvolNet is to optimize the weights of LightdehazeTestnet to improve the Natural Image Quality Evaluator (NIQE), Entropy, and Structural Similarity Index Measure (SSIM) scores. The method used in this work can be seen on Figure 1 and consists of the following steps. The first step involves creating a starting population for the evolutionary algorithm. This is achieved by training the CNN using image enhancement algorithms, Unsharp Masking (UM), High-Frequency Emphasis Filtering (HEF), Contrast Limited Adaptive Histogram Equalization (CLAHE), Automatic Tissue Attenuation-based Contrast Enhancement (ATACE), and Triple Clipped Dynamic Histogram Equalization (TCDHE), as references and using weights at different epochs for the starting population. Subsequently, the starting population undergoes optimization through selection, permutation, and mutation. Selection is based on a fitness function that ranks different weights using NIQE, entropy, and SSIM scores, with the best weights selected as parents for the next generation. Permutation and mutation are then applied to these selected weights to produce the next generation. At the conclusion of the evolutionary algorithm, the best performing weight is chosen as the optimal solution. All experiments were done using the Google Colab Python 3 environment with a T4 GPU and a high RAM option.

3.1 Statistical metrics evaluation

To assess the image quality of the chest X-ray images, three metrics were employed: Natural Image Quality Evaluator (NIQE), entropy, and Structural Similarity Index Measure (SSIM). NIQE and entropy were utilized for no-reference image quality assessment (IQA), while SSIM was employed to measure the distortion between the original and enhanced image.

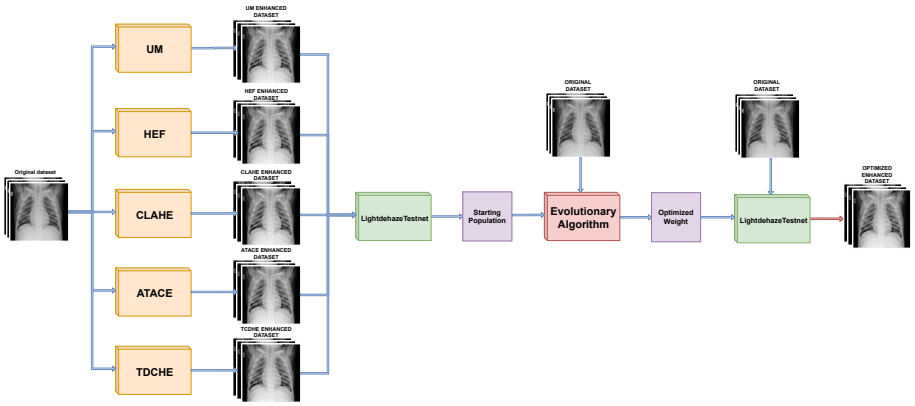

Figure 1 – IQAEvolNet methodology diagram.

3.1.1 Entropy

Entropy is a mathematical function based on the distribution of different pixel values in the image histogram. The formula is expressed in Equation 3.1, where p_i is the value of the normalized histogram for the pixel intensity i. As the Entropy is a measure of how evenly distributed the histogram is, the best-case scenario consists of a flat histogram where all pixel levels have the same value. This means that, for an image with the typical 0 to 255 pixel value, p_i would be equal to $256^{-1} = 2^{-8}$. By applying this value to Equation 3.1 we have that the maximum value for the Entropy is 8.

$$Entropy = -\sum_{0}^{n-1} p_i \log_2 p_i \tag{3.1}$$

The worst-case scenario would be where all pixels have the same value. In this case, p_i would be equal to either 1 or 0 and, therefore, either $p_i = 0$ or $\log_2 p_i = 0$. This means that all the sum terms of Equation 3.1 would be 0 and the Entropy value would also be 0. In conclusion, Entropy is a no-reference IQA metric that measures the distribution of the histogram of an image. The higher the metric, the more evenly distributed the histogram. All values fall within the range of 0 to 8.

3.1.2 NIQE

NIQE is a no-reference IQA algorithm, proposed by (MITTAL *et al.*, 2013), based on the statistical properties of images. Unlike other methods, NIQE does not rely on datasets with human-graded images for training. The model comprises a collection of quality-aware features that are fitted into a multivariate Gaussian model. The quality of an image is then determined

by the distance of its Gaussian model from the average Gaussian model of all natural images. The higher the value, the more distant the image is from the average and therefore the lower its assumed quality. In contrast to entropy, where higher values indicate better quality, with NIQE, higher values suggest worse image quality. The range of values has no fixed upper limit, but due to its distance logic, there is a lower limit at 0. Therefore, all NIQE values fall within the $[0, +\infty[$ range.

3.1.3 SSIM

The no-reference IQA metrics, in their own right, do not provide a complete picture. Since they operate without relying on the original image, they do not consider the preservation of information. Therefore, while an enhanced image may have higher quality than the original according to these metrics, it must also conserve information. Relying solely on no-reference IQA metrics could lead to a scenario where an algorithm generates random high-quality images, despite the output having no relation to the input.

To address this issue, similarity metrics were incorporated into the fitness function. This ensures that the generated solutions not only improve the no-reference IQA metrics but also minimize the changes made to the original image to achieve this improvement. The chosen similarity metric is SSIM, proposed by (WANG *et al.*, 2004) as a means to measure the distortion between a distorted image and a reference image. SSIM depends on three different characteristics: luminance, contrast, and structure. Luminance comparison is done using Equation 3.2. Luminance comparison is performed using Equation 3.2. The constant C_1 is introduced to prevent steep increases when $\mu_x^2 + \mu_y^2$ is close to zero. The value of C_1 is defined in 3.3 where K_1 is a small constant and L represents the range of pixel values, which is 255 in this case.

$$l(x,y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \tag{3.2}$$

$$C_1 = (K_1 L)^2 \tag{3.3}$$

The contrast comparison is conducted using a similar formula, with the mean values replaced by the standard deviations. Equation 3.4 presents the formula. The constant C_2 serves a similar purpose to C_1.

$$c(x,y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \tag{3.4}$$

$$C_2 = (K_2 L)^2 \tag{3.5}$$

The structure comparison is performed using a similar formula, but instead of using mean and standard deviation, it involves the correlation coefficient between the normalized images. Equation 3.6 shows the formula. The constant C_3 serves a similar purpose to C_1 and C_2.

$$s(x,y) = \frac{2\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \tag{3.6}$$

$$C_3 = \frac{C_2}{2} \tag{3.7}$$

Finally, SSIM is calculated by multiplying all three comparisons. Equation 3.8 shows the final formula.

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{3.8}$$

Since all three comparison metrics have values in the interval [0,1], the product of these metrics, SSIM, also has values in the interval [0,1].

3.2 LightdehazeTestnet

The CNN model employed in this study, LightdehazeTestnet, is derived from the Lightdehazenet architecture proposed in (ULLAH *et al.*, 2021). While the general structure of the model was retained, modifications were made to the output image equation, and an additional convolutional layer was introduced. In the version used for this work, the model calculates two correction maps: a correction coefficient C(x) and a correction bias B(x). The correction coefficient C(x) is applied to the input image through multiplication, while the correction bias B(x) is added for fine adjustment.

Figure 2 illustrates the architecture of the model. ReLu and Sigmoid represent the Rectilinear Uniform and Sigmoid activation layers, respectively. Batch Norm stands for batch

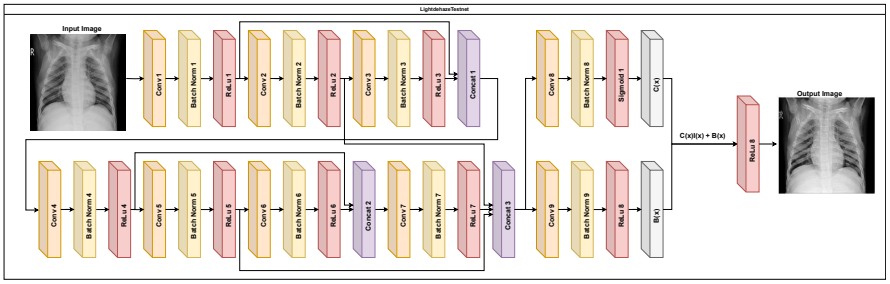

Figure 2 – LightdehazeTestnet model architecture.

Table 1 – Parameters for the Convolutional layers of the LightdehazeTestnet model.

Layer	Kernel size	Stride	Input channels	Output channels
Conv 1	1	1	1	8
Conv 2	3	1	8	8
Conv 3	5	1	8	8
Conv 4	7	1	16	16
Conv 5	3	1	16	16
Conv 6	3	1	16	16
Conv 7	3	1	32	32
Conv 8	3	1	56	1
Conv 9	3	1	56	1

normalization layers, Concat denotes concatenation layers, and Conv refers to convolutional layers. The model employs convolutional operations to generate feature maps with dimensions identical to those of the input image. In order to amplify the influence of the deepest layers on the output, previous feature maps are reintroduced into the model through concatenation.

The parameters utilized in the convolutional operations are detailed in Table 1. To ensure that all feature maps match the dimensions of the input image, the stride is consistently maintained at 1. The size of the kernels varies to capture different ranges of information. Concatenation is achieved by simply appending the channels of the feature maps.

3.3 IQAEvolNet

The solution proposed by this work relies on the random mixing and mutation of existing solutions. Therefore, before discussing the evolutionary algorithm itself, it's important

to explain the generation process of the starting solutions used in the experiments.

3.3.1 Creation of the starting population

The first step of IQAEvolNet involves creating the initial population of weights. This process was carried out by training the LightdehazeTestnet model using the reference image enhancement algorithms. Initially, an enhanced reference dataset was generated using each image enhancement algorithm. Subsequently, back-propagation was employed to train the model by comparing its results with the reference dataset. At the conclusion of each training epoch, the weights of the model were saved. These weights served as the starting population.

For the creation of the dataset for the initial training of the model, the following image enhancement algorithms were utilized:

- UM with a radius of size 5 and a scaling factor of 2;
- HEF with D_0 of 20, K_1 of 0.5 and K_2 of 0.75;
- CLAHE with parameters (8,8) for the tile grid size and a clipping limit of 2.0;
- ATACE with a local region of size (7,7);
- TCDHE with a β of 0.5.

For the back-propagation training, the parameters were set as follows:

- Number of epochs: 100;
- Optimizer: Adam;
- Learning rate: 0.0001;
- Learning rate decay: 0.0001;
- Loss function: Mean Square Error.

Weights from epochs 60, 70, 80, 90, and 100 were selected from each of the five algorithms for the starting population, resulting in a total of 25 (5×5) individuals. The next step of the algorithm involves defining the fitness function to be used.

3.3.2 Fitness Function

The fitness function utilized by the algorithm incorporates two IQA metrics and one similarity metric: Entropy, NIQE, and SSIM. These metrics are employed to evaluate the effectiveness of the enhancement models.

Equation 3.9 shows the formula for the fitness function. The weights w_1, w_2, and w_3 are used to calibrate the priority between the metrics. The minus sign on w_2 was done because

NIQE, contrary to the other metrics, improves the lower its value.

$$Fitness = Entropy^{w_1} \times NIQE^{-w_2} \times SSIM^{w_3} \tag{3.9}$$

3.3.3 Evolutionary algorithm

The evolutionary algorithm operates by evaluating the fitness of a set of solutions, selecting the best-performing ones, and generating a new set of solutions from them. This process iterates using the newly generated set of solutions as the starting population for the next iteration.

Figure 3 depicts the operation of the algorithm. The population and individuals refer to the weights for the LightdehazeTestnet. The algorithm operates in batches. At the beginning of each batch, each weight set that composes the starting population is loaded into the LightdehazeTestnet model and applied to the batch of images. The resulting enhanced batches undergo a fitness test, where each receives a score quantifying its image quality. This fitness test ranks the individuals of the starting population, preserving the highest-scoring individuals as the surviving population and discarding the rest. The best-performing individual is designated as the optimal individual. The surviving population then undergoes permutation and mutation processes to create a new population with the same size as the starting population. This new population serves as the starting population for the next batch. The algorithm continues to run, selecting new images as samples, until the dataset is exhausted. At this juncture, it is deemed that one epoch or generation of training has been completed.

If it is the first epoch of training, the best performing individual in the last batch is designated as the solution for that epoch. Otherwise, the fitness of this individual is compared with that of the previously best performing individual across the entire dataset. If it exhibits higher fitness, it is deemed the new best performing individual; otherwise, the algorithm determines that this epoch did not produce improvement. To make this assessment, a threshold is employed when comparing the fitness of both individuals. The new individual must surpass the previous one by at least a specified proportion to be considered an improvement.

The fitness test is conducted using the fitness function presented previously. The metric values utilized in the equation are the averages of the dataset sample. The algorithm arranges the population by fitness in descending order and selects the square root of n best performing individuals, where n is the size of the starting population.

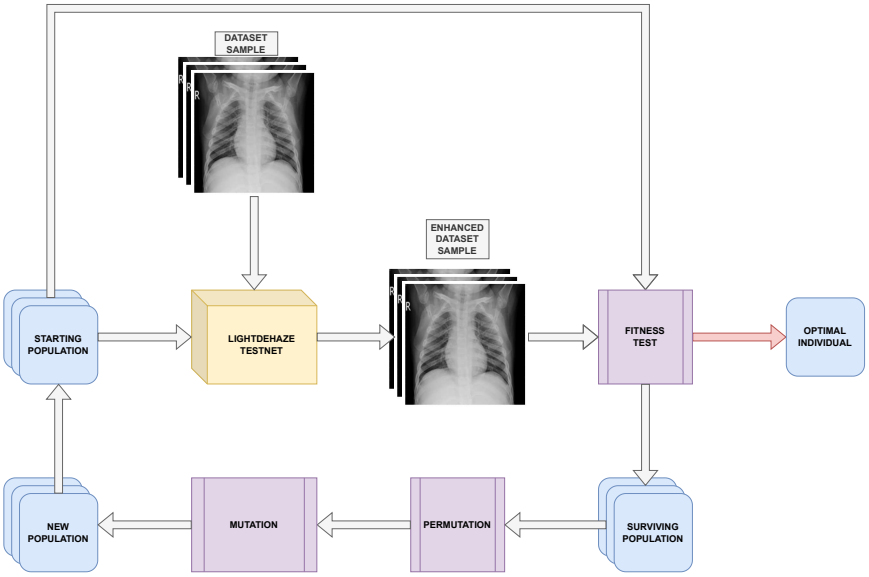

Figure 3 – IQAEvolNet algorithm diagram. Yellow square represent LightdehazeTestnet model, blue squares represent LightdehazeTestnet weights and purple squares represent evolutionary processes.

The permutation operation involves creating a new set of solutions by combining the surviving population. This new set of solutions comprises the square root of n surviving individuals and their $\sqrt{n}(\sqrt{n}-1)$ paired combinations, which collectively return us to the population size of n. In the implementation of this work, the paired combination was achieved by calculating the weighted average between the two solutions, with the first weight being a random number following a Gaussian distribution with mean = 0.5 and standard deviation = $\frac{0.5}{3}$. These solutions are then subjected to mutation.

The mutation operation, as depicted in Equation 3.10, entails multiplying the starting parameter by a random variable following a normal distribution with a mean of 1 and an adjustable standard deviation called the noise intensity.

$$M(x,y) = I(x,y) \cdot N(1,\sigma) \tag{3.10}$$

In this case, the solution has two types of parameters: convolutional filters and bias vectors. In this implementation, if a mutation occurs, it will affect all elements of a specific convolutional filter or bias vector. For example, consider a CNN with only one convolutional layer featuring a 3x3 kernel with 8 output channels and a bias. This implies that each of the 8 kernels and the CNN bias vector will have the potential to trigger a mutation. When activated, the mutation will involve multiplying the elements of the kernel or bias vector by a matrix or vector of the same dimensions, where the terms of this matrix or vector follow a normal distribution.

Table 2 – Main Dataset class composition and IQA metrics mean and standard deviation. NIQE is better at lower values and entropy is better at higher values.

Class	Number of images	NIQE ↓	Entropy ↑
Viral pneumonia	1345	5.76(0.79)	7.210(0.244)
Normal	10192	5.76(0.89)	7.482(0.319)
Lung Opacity	6012	6.00(1.25)	7.377(0.362)
COVID	3616	6.20(1.51)	7.272(0.385)
Whole dataset	21165	5.92(1.16)	7.388(0.354)

Once the mutation process is completed, the new population replaces the old one, and the algorithm determines whether to continue the process or not. If the algorithm is still in the middle of an epoch or the stopping criteria haven't been reached, it iterates again; otherwise, it stops and returns the current best-performing solution.

Finally, the noise intensity is reduced each epoch the algorithm fails to find a better solution. Each time this happens, the noise intensity is multiplied by an adjustable factor called noise decay.

For the implementation used in this work, the population size was set to $n = 25$, the improvement threshold was 1.00001, the batch size was 32, the mutation probability was 0.1, the initial noise intensity was 1, and the noise decay factor was 0.5.

3.4 Datasets

The main dataset utilized in this study,(CHOWDHURY *et al.*, 2020) (RAHMAN *et al.*, 2021b), comprises an aggregation of multiple chest X-rays. It consists of 1345 viral pneumonia, 10192 normal, 6012 lung opacity, and 3616 Covid chest X-ray images. Figure 4 displays an example of each class. For the experiments, all images were resized to 224 $\times$ 224 pixels using bi-cubic interpolation. The dataset has an average NIQE value of 5.92 with a standard deviation of 1.16, and an average entropy value of 7.388 with a standard deviation of 0.354. Table 2 provides the number of images and the mean and standard deviation of image quality assessment metrics for each class and the entire dataset.

Two extra datasets were also used for the image classification experiments: Shenzen dataset, (JAEGER *et al.*, 2014), and NIH Chest X-ray dataset, (WANG *et al.*, 2017). The Shenzen dataset consists of 336 Tuberculosis and 326 Normal images. The NIH Chest X-ray dataset consists of more than 100000 images with multiple possible diagnostics. For simplicity the dataset used in the experiment was a filtered version with 2 possible diagnostics: Filtered and

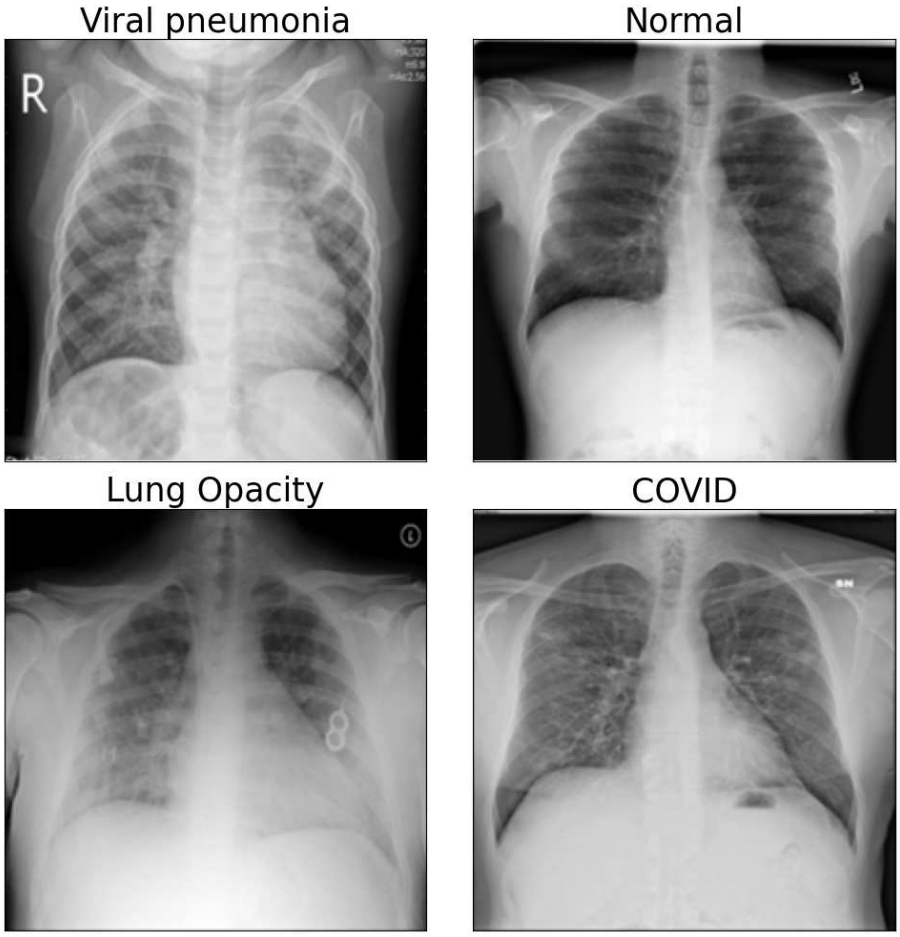

Figure 4 – Image example for each class of the main dataset.: (**a**) Viral Pneumonia. (**b**) Normal. (**c**) Lung Opacity. (**d**) COVID.

Normal. The filter was done prioritizing low quality images to test whether the enhancement algorithms impact on classification are more notable in low quality datasets. First, a quality score was calculated for each image by dividing the entropy score by the NIQE score. Then, the 20000 images with the lowest score were selected. Finally, the two most common diagnostic were preserved while the rest was excluded. Finally, the number of images per diagnosis was limited to the lowest scoring 4000 to better balance the classes. The result was a filtered dataset with 4000 Normal and 2307 Infiltration diagnosis. For the experiments, all images were resized to 224 $\times$ 224 pixels using bi-cubic interpolation.

3.5 Classification test

To further validate the proposed method an experiment was performed to compare the impact of the algorithms on a classification task. The classification was done using transfer learning with the Mobilenet, (HOWARD *et al.*, 2017), model as a base. The chosen top layer

consisted of a flatten layer, followed by a dense layer with 128 units and a rectified linear unit activation function, then a dropout layer with a 0.3 rate, and finally a dense layer, with 4 units for the main dataset and 2 units for the Shenzen and NIH datasets, and a softmax activation function. The training was done in two stages, an initial stage and a fine-tuning stage. In both stages, the optimizer utilized was stochastic gradient descent with a 0.9 momentum. In the first stage the model was trained with a learning rate of 10^{-4}, the maximum number of epochs was set to 20 and an early-stopping of 5 epochs without improvement. In the fine tuning stage the learning rate was set to 10^{-6}, the maximum number of epochs was set to 10 and an early-stopping of 5 epochs without improvement. The dataset was split into training and validation subsets with an 80%/20% ratio respectively. The loss function chosen was Categorical Cross-entropy.

4 RESULTS AND DISCUSSION

In this section it is discussed the experimental results of the work. First, in section 4.1, it is analyzed the effectiveness of the non-learning image enhancement algorithms using Structural Similarity Index Measure (SSIM), Natural Image Quality Evaluator (NIQE), entropy, and processing time. Then, in section 4.2, it is presented the results for the LightdehazeTestnet trained using each of the non-learning image enhancement algorithms as reference. Following this, in section 4.3, it is examined the results for optimizing the weights of LightdehazeTestnet using only NIQE and entropy, without relying on SSIM. In sequence, in section 4.4, it is discussed the results for optimizing the weights of LightdehazeTestnet using all three metrics. Furthermore, in section 4.5 the impact of the enhancement algorithm in the classification task is analyzed. Finally, in section 4.6 it is studied the metric changes across epochs during the evolutionary algorithm process and the impact of changing the weights of the fitness function in the final result.

4.1 Non-learning image enhancement

First, the image quality assessment (IQA) metrics for the non-learning image enhancement methods were calculated to be used as a reference. Table 3 shows the results. The numbers shown are the mean and standard deviation for the SSIM, NIQE, entropy, and processing time of the methods across the entire dataset.

The method with the best SSIM and processing time is UM, while also exhibiting the worst NIQE and entropy. This suggests that UM makes the least modifications to the original images but also produces the lowest-quality enhancements. HEF, on the other hand, achieves the best NIQE and entropy scores but has the slowest processing time and the second-worst SSIM. Despite being the most effective in terms of image quality, HEF introduces considerable distortion and requires more processing time. CLAHE demonstrates the second-best NIQE result and processing time, along with a moderate SSIM and the second-worst entropy. It stands out for its processing efficiency and preservation of image content. ATACE exhibits the poorest SSIM and moderate scores in NIQE, CLAHE, and processing time. It appears to induce the most significant alterations from the original image. TCDHE achieves the second-best SSIM and entropy scores, with the second-worst NIQE and processing time. It manages to attain high entropy while making minimal changes to the original image. In conclusion, HEF emerges as the

Table 3 – Reference image enhancement algorithms IQA metrics results for the entire dataset. ↑ means that higher values are better, ↓ means the opposite. Bold numbers highlight the best results.

Enhancement Method	SSIM ↑	NIQE ↓	Entropy ↑	Time (seconds) ↓
UM	**0.970(0.008)**	10.09(3.24)	7.409(0.351)	**4.64E-04(3.07E-04)**
HEF	0.729(0.102)	**4.24(0.70)**	**7.994(8.05E-08)**	1.95E-01(5.68E-03)
CLAHE	0.849(0.032)	4.99(0.74)	7.536(0.295)	5.69E-04(4.42E-04)
ATACE	0.570(0.103)	5.53(1.35)	7.589(0.304)	1.89E-03(2.78E-04)
TCDHE	0.958(0.039)	5.89(1.08)	7.665(0.262)	1.18E-01(1.72E-02)

Table 4 – IQA metrics results for Figure 5. ↑ means that higher values are better, ↓ means the opposite. Bold numbers highlight the best results.

Enhancement Method	SSIM ↑	NIQE ↓	Entropy ↑
Original image	-	5.47	7.399
UM	**0.955**	7.88	7.438
HEF	0.747	**4.18**	**7.994**
CLAHE	0.835	4.64	7.573
ATACE	0.597	5.48	7.437
TCDHE	0.945	4.76	7.746

best enhancement method in terms of image quality, although it comes with higher processing time.

Figure 5 shows the results for each image enhancement method and Table 4 shows the IQA metrics for each image. It is evident that HEF achieves the highest level of quality among all enhancement methods, as expected from the entropy and NIQE values. Conversely, ATACE produces the most distinct enhanced images from the originals, as indicated by the SSIM values. In contrast to this, the UM method generates images very similar to the originals, with only a slight increase in edge prominence, consistent with its high SSIM. Its elevated NIQE value might suggest a significant drop in quality, which isn't observed in practice. This example shows that NIQE is not a perfect representation of chest X-ray image quality. CLAHE, as expected from its SSIM and NIQE results, presents a more modest contrast enhancement compared to HEF but yields enhanced images more akin to the originals. TCDHE exhibits images very close to the originals, consistent with its high SSIM value. While a slight increase in contrast is discernible, it's less pronounced than in other methods except UM. This deviation from expectations, especially given UM's higher NIQE value but lower entropy, indicates that entropy isn't a perfect gauge of chest radiography image quality either.

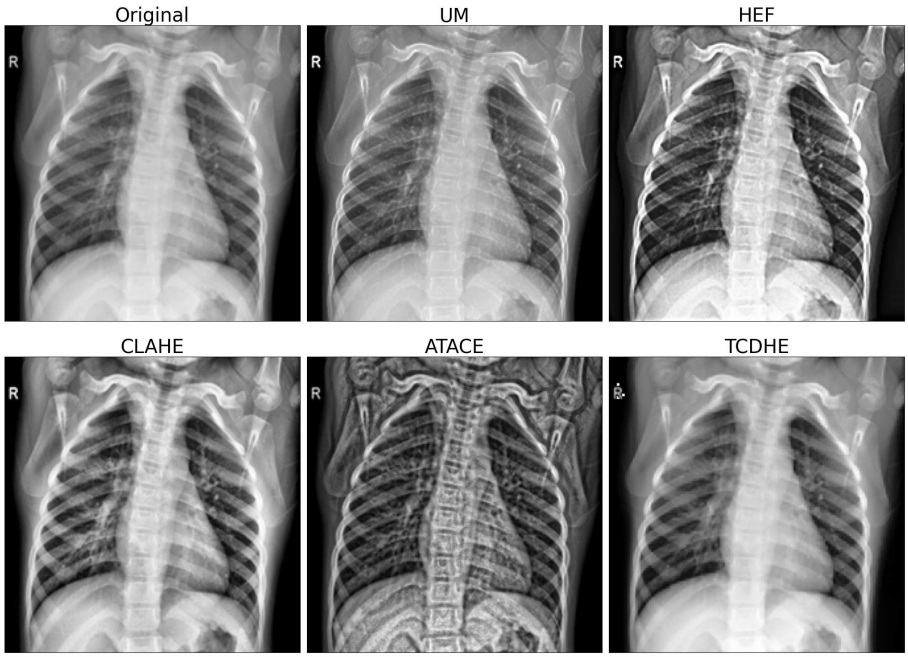

Figure 5 – Image enhancement example with the literature methods: (**a**) Original image. (**b**) Unsharp Masking. (**c**) High-Frequency Emphasis Filtering. (**d**) Contrast Limited Adaptive Histogram Equalization. (**e**) Automatic Tissue Attenuation-based Contrast Enhancement. (**f**) Triple Clipped Dynamic Histogram Equalization.

4.2 Pre-trained lightdehazeTestnet

In this section, the results from pre-training the LightdehazeTestnet using each of the algorithms as reference are explored. Table 5 presents the IQA metrics and processing time for these experiments. The average processing time remained consistent across the experiments, with most achieving an average processing time of 2.90 milliseconds. Comparing these results with Table 3, it is evident that the CNN pre-trained using UM exhibited improved NIQE, from 10.09 to 5.09, and better Entropy, from 7.409 to 7.447, compared to UM itself. Although the SSIM decreased slightly, from 0.970 to 0.937, it still remained relatively high compared to other methods. For the CNN pre-trained using HEF, the average NIQE changed minimally, from 4.24 to 4.26, while the SSIM and Entropy worsened, from 0.729 to 0.686, and from 7.994 to 7.798, respectively. However, there was a significant reduction in processing time, from 0.195 seconds to 2.90 milliseconds, nearly a hundredfold decrease. The CNN pre-trained using CLAHE showed an improvement in average NIQE, from 4.99 to 4.53, a slight reduction in Entropy, from 7.536 to 7.522, and a notable decrease in SSIM, from 0.849 to 0.798. For the CNN pre-trained using ATACE, the NIQE and SSIM remained similar, from 0.570 to 0.573, and from 5.53 to 5.56, respectively, while there was a minor improvement in entropy, from 7.589 to 7.610. The CNN pre-trained using TCDHE exhibited a slight worsening of SSIM, NIQE, and Entropy, from 0.958

Table 5 – IQA metrics and processing time for pre-trained lightdehazeNet using different algorithms as reference. ↑ means that higher values are better, ↓ means the opposite. Bold numbers highlight the best results.

Enhancement method	SSIM ↑	NIQE ↓	Entropy ↑	Time (seconds) ↓
Pre-trained CNN with UM	**0.937(0.093)**	5.09(1.22)	7.447(0.266)	2.94E-03(6.30E-03)
Pre-trained CNN with HEF	0.686(0.086)	**4.26(0.67)**	**7.798(0.173)**	**2.90E-03(1.21E-04)**
Pre-trained CNN with CLAHE	0.798(0.071)	4.53(0.62)	7.522(0.209)	**2.90E-03(1.29E-04)**
Pre-trained CNN with ATACE	0.573(0.071)	5.56(1.22)	7.610(0.234)	**2.90E-03(1.89E-04)**
Pre-trained CNN with TCDHE	0.929(0.078)	5.81(1.08)	7.517(0.235)	**2.90E-03(1.17E-04)**

to 0.929, from 5.89 to 5.81, and from 7.665 to 7.517, respectively. However, there was a notable improvement in processing time, from 0.118 seconds to 2.90 milliseconds. This experiment demonstrates that pre-trained CNNs can achieve comparable results to the reference algorithms, and for HEF and TCDHE, it also significantly improves processing time.

Figure 6 depicts the results of applying the pre-trained CNNs. Similar to Figure 5, the results display enhanced images. However, two of the displayed images exhibit noticeable issues. In the pre-trained CNN with HEF example, black dots are clustered around the bottom of the image, particularly in regions with a high concentration of pixels with elevated values. In the pre-trained CNN with CLAHE example, a deformed halo artifact is visible in the same area. These artifacts highlight that the LightdehazeTestnet CNN used is not a perfect fit and can be further optimized for chest X-ray image enhancement. However, the primary focus of this work is not on the adequacy of LightdehazeTestnet for this application but rather on demonstrating that an evolutionary approach can effectively train a CNN to optimize NIQE and Entropy scores. These issues may also be present in other results, but the primary goal of this work is not to provide an ideal CNN for chest X-ray image enhancement, but to illustrate that the evolutionary method can successfully train a CNN for optimizing NIQE and Entropy scores.

4.3 IQAEvolNet with only IQA metrics

The initial evolutionary test was conducted using only IQA metrics as fitness indicators, excluding SSIM. The results of this experiment are presented in Table 6. Weight set 1 was trained using only NIQE as a fitness measure. As anticipated, this weight set achieved the lowest NIQE score. It also attained the highest SSIM score. However, it is noteworthy that the achieved entropy value of 6.924 is lower than the entropy of the original dataset, which was 7.389. Weight set 2 was trained using only Entropy as a fitness measure. Although this weight set achieved the highest entropy score, it exhibited a significantly low SSIM, indicating substantial image

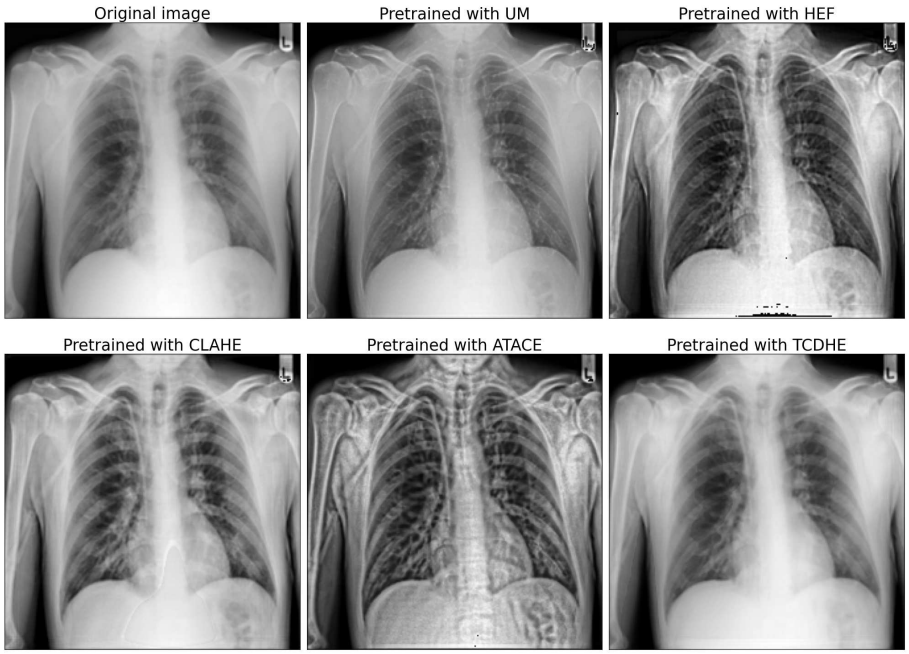

Figure 6 – Experimental results for the lightdehazeNet pre-trained using each of the enhanced algorithms: (**a**) Original image. (**b**) Pre-trained with UM. (**c**) Pre-trained with HEF. (**d**) Pre-trained with CLAHE. (**e**) Pre-trained with ATACE. (**f**) Pre-trained with TCDHE.

Table 6 – Results for evolutionary training without SSIM. ↑ means that higher values are better, ↓ means the opposite. Bold numbers highlight the best results.

Weight set	SSIM weight	NIQE weight	Entropy weight	SSIM ↑	NIQE ↓	Entropy ↑
Weight set 1	0	1	0	**0.525(0.097)**	**3.20(0.74)**	6.924(0.252)
Weight set 2	0	0	1	0.042(0.014)	19.30(3.75)	**7.950(0.136)**
Weight set 3	0	1	1	0.495(0.085)	3.71(0.71)	7.763(0.171)

distortion. Additionally, it obtained a very high NIQE score of 19.30 compared to the original dataset's NIQE of 5.62, suggesting a resulting image of very low quality. Weight set 3 was trained using both NIQE and Entropy as fitness measures, with equal weights. It demonstrated improvements in both NIQE and entropy compared to the original dataset, with NIQE decreasing from 5.92 to 3.71 and entropy increasing from 7.389 to 7.763.

Figure 7 illustrates an image enhancement example for the models. Weight sets 1 and 2, which utilize only one IQA metric for fitness calculation, generate highly distorted images. In weight set 3, where both IQA metrics are used, the resulting image exhibits higher quality. However, a notable issue observed in this image is the presence of a large cluster of black dots in the region with a high concentration of high-value pixels.

This result underscores the limitation of solely relying on IQA metrics for fitness calculation, as it leads to models that distort the original image. Without a reference point, these metrics cannot fully capture the information lost during the enhancement process. To address

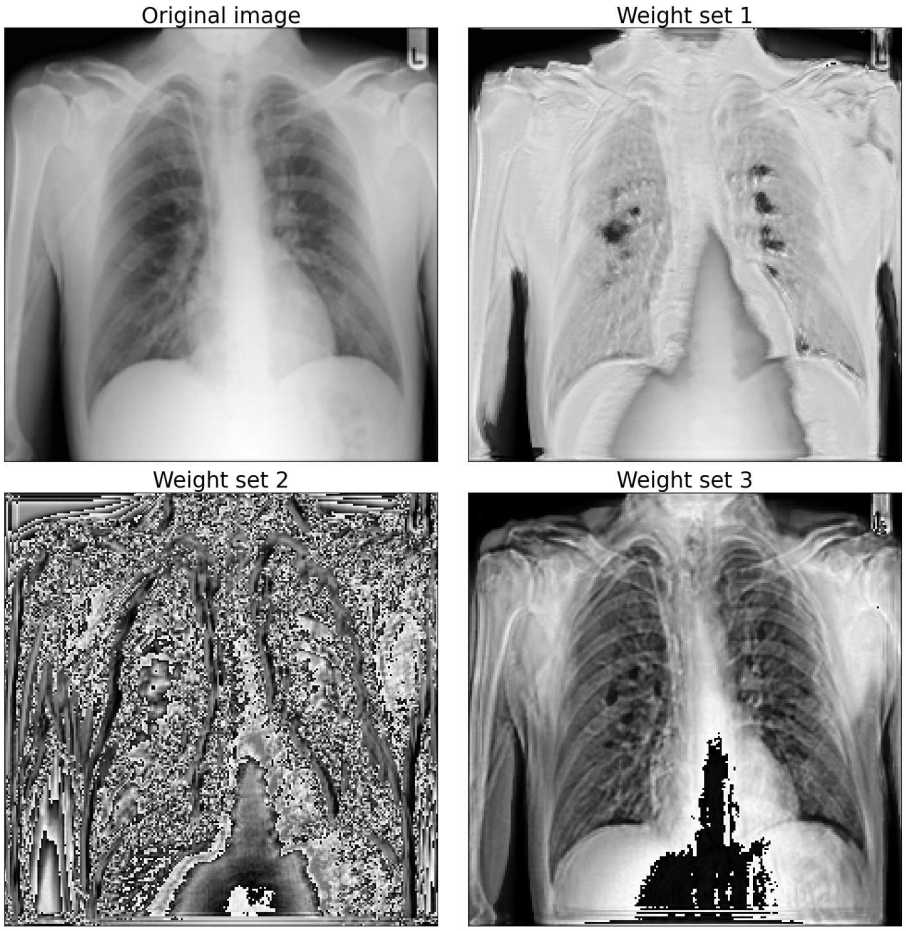

Figure 7 – Image enhancement example for the models trained without SSIM: (**a**) Original image. (**b**) Weight set 1. (**c**) Weight set 2. (**d**) Weight set 3.

this issue, SSIM is incorporated into the fitness calculation. This ensures that the generated models must strike a balance between increasing the IQA metric and minimizing the distortions they introduce.

4.4 IQAEvolNet with IQA metrics and SSIM

In this subsection, the results obtained using different weight sets for fitness calculation were presented. The average SSIM, NIQE, and entropy for each weight set are presented in Table 7. The two best-performing weight sets from an SSIM standpoint were weight set 5 and 9, with SSIM values of 0.848 and 0.851, respectively. This outcome was expected since they have the highest SSIM weight. Both weight sets exhibit good NIQE values, 3.88 and 4.05 respectively, compared to the best-performing algorithm, HEF, with a NIQE of 4.18. However, weight set 5 has a low entropy value of 7.251, even lower than the original dataset, which has a value of 7.389. In contrast, weight set 9 has a competitive entropy value of 7.576, higher than the scores for UM (7.409) and CLAHE (7.536), and only slightly lower than ATACE (7.589).

Therefore, it is concluded that weight set 9 yields good results with minimal distortion.

The best-performing weight set from a NIQE perspective was weight set 7, which had the highest NIQE weight. Among all the state-of-the-art algorithms analyzed in this work, the best-performing one from a NIQE perspective was HEF, which achieved an average NIQE score of 4.18. The evolved model managed to achieve an average NIQE score of 3.48, representing a significant improvement. However, this model only managed to achieve an entropy score of 7.503, which is very similar to the original dataset score of 7.399 and worse than HEF (7.994), CLAHE (7.573), and TCDHE (7.746). Due to this under-performance in entropy, this weight set is not considered a good enough solution.

The best-performing weight set from an entropy point of view was weight set 11. It achieved a mean entropy of 7.799, a significant improvement from the original dataset score of 7.399 and bigger than all state-of-the-art algorithms studied in this work. It simultaneously achieve an average NIQE of 3.620, bigger than the NIQE score of all reference algorithms, and close to the best performing weight set NIQE-wise of 3.48. From a combined analysis of NIQE and entropy, this weight set was the best-performing from all the tested weight sets. But it's low SSIM value of 0.645, lower than all reference algorithms except ATACE, suggests that it's a model prone to creating distorted images.

The best-performing weight set from an entropy perspective was weight set 11. It achieved a mean entropy of 7.799, a significant improvement from the original dataset score of 7.399 and higher than all state-of-the-art algorithms studied in this work, except HEF. Simultaneously, it achieved an average NIQE of 3.620, higher than the NIQE score of all reference algorithms and close to the best-performing weight set NIQE-wise of 3.48. From a combined analysis of NIQE and entropy, this weight set was the best-performing among all the tested weight sets. However, its low SSIM value of 0.645, lower than all reference algorithms except ATACE, suggests that it's a model prone to creating distorted images.

In Figure 8, an example of image enhancement using the models trained for each weight set is shown. Table 8 displays the IQA metric scores for these images. Among the weight sets, 9 and 11 were considered good solutions based on their IQA score average value. Weight set 11 generates a notable amount of black dot clusters, which was expected given its low SSIM score. Although it has the best entropy score (7.821), it also has the worst NIQE and SSIM, indicating significant distortion. Therefore, weight set 11 did not generate a satisfactory solution in this example. Weight set 9 also generates small clusters of black dots but in a

Table 7 – IQA metric results for each weight set. ↑ means that higher values are better, ↓ means the opposite. Bold numbers highlight the best results.

Weight set	SSIM weight	NIQE weight	Entropy weight	SSIM ↑	NIQE ↓	Entropy ↑
Weight set 4	1	1	1	0.795(0.066)	3.55(0.78)	7.473(0.250)
Weight set 5	2	1	1	0.848(0.058)	3.88(0.89)	7.251(0.303)
Weight set 6	1	1	2	0.800(0.071)	3.95(0.94)	7.550(0.291)
Weight set 7	1	2	1	0.708(0.100)	**3.48(0.87)**	7.503(0.248)
Weight set 8	1	1	4	0.838(0.077)	4.11(1.05)	7.544(0.264)
Weight set 9	2	1	4	**0.851(0.086)**	4.05(1.02)	7.576(0.244)
Weight set 10	1	1	8	0.650(0.091)	4.01(0.67)	7.763(0.194)
Weight set 11	1	1	16	0.645(0.071)	3.62(0.60)	**7.799(0.187)**

Table 8 – IQA metric score for Figure 8. ↑ means that higher values are better, ↓ means the opposite. Bold numbers highlight the best results.

Weight set	SSIM ↑	NIQE ↓	Entropy ↑
Original image	-	5.02	7.402
Weight set 4	0.788	**2.39**	7.443
Weight set 5	**0.887**	2.88	7.164
Weight set 6	0.831	3.78	7.533
Weight set 7	0.750	3.35	7.533
Weight set 8	0.813	2.91	7.500
Weight set 9	0.869	3.09	7.547
Weight set 10	0.683	3.36	7.713
Weight set 11	0.628	4.06	**7.821**

significantly smaller amount compared to weight set 11. It achieved the second-highest SSIM, the fourth-lowest NIQE, and the fifth-highest entropy. Despite other images struggling with low contrast or excessive black dot distortion, weight set 9 maintained a balance between both IQA metrics while keeping SSIM high. Therefore, weight set 9 was identified as the optimal solution for this image example. The fact that weight set 11 was not chosen as the optimal solution despite outperforming both IQA metrics suggests that these metrics do not accurately quantify the image quality of chest X-ray images.

4.5 Classification results

The reference algorithms and the proposed methods were used as pre-processing for the classification of the main dataset. The results are shown in Table 9. It is observed that all enhancement algorithms had little effect on the classification accuracy. HEF and UM, with accuracy of 90.56% and 90.87% respectively, had a smaller accuracy than with no pre-processing algorithms (91.19%). The other algorithms showed an slight increase in accuracy, with the biggest increase being TCDHE which achieved 91.43% of accuracy. Similar results

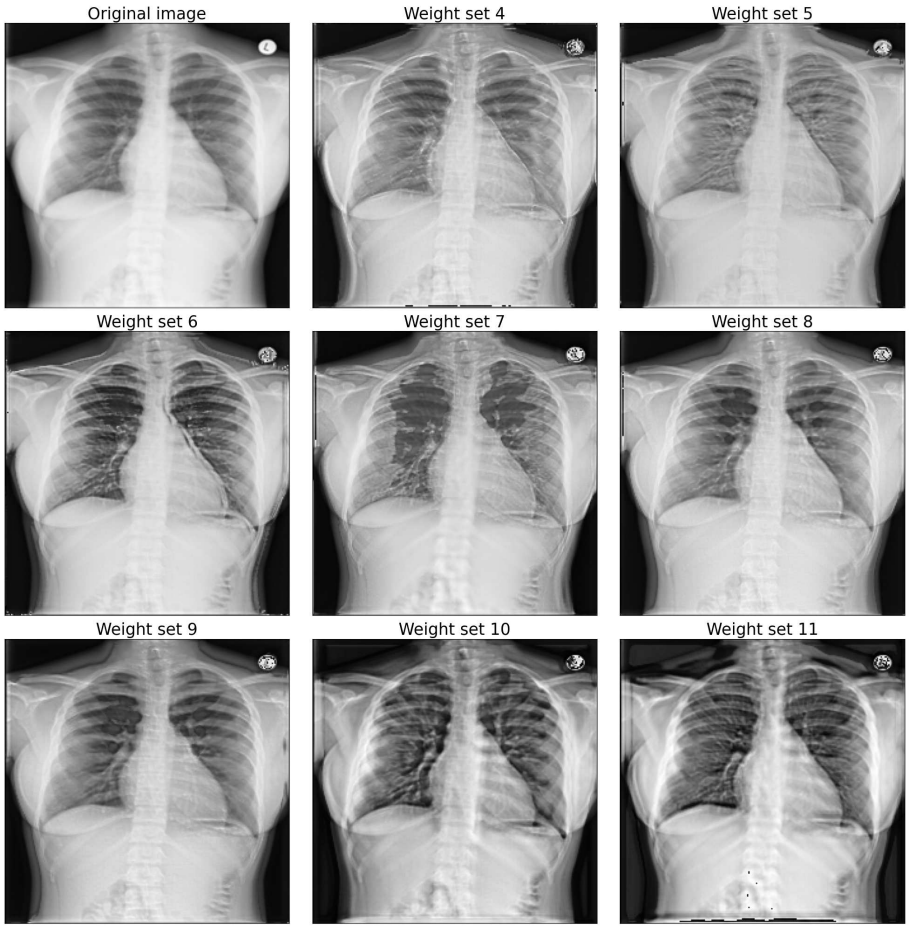

Figure 8 – Image enhancement example for the models trained SSIM: (**a**) Original image. (**b**) Weight set 4. (**c**) Weight set 5. (**d**) Weight set 6. (**e**) Weight set 7. (**f**) Weight set 8. (**g**) Weight set 9. (**h**) Weight set 10. (**i**) Weight set 11.

Table 9 – Classification results for each of the image enhancement algorithms. Bold numbers highlight the best results.

Algorithm	Accuracy
None	91.19%
UM	90.87%
HEF	90.56%
CLAHE	91.38%
ATACE	91.24%
TCDHE	**91.43%**
IQAEvelNet	91.35%

were achieved on the Shenzen and NIH datasets as shown in Tables 10 and 11. In all cases, the changes in accuracy were minimal. This goes against the results presented in (RAHMAN *et al.*, 2021a), where CLAHE increased the accuracy of classification algorithms for COVID-19 diagnosis, and (MUNADI *et al.*, 2020), where UM, HEF and CLAHE increased the accuracy of classification algorithms for tuberculosis detection up to 3.37%.

The results show that image enhancement algorithms by themselves have very little

Table 10 – Classification results for each of the image enhancement algorithms on the Shenzen dataset. Bold numbers highlight the best results.

Algorithm	Accuracy
None	87.31%
UM	83.58%
HEF	87.31%
CLAHE	**89.55%**
ATACE	85.82%
TCDHE	85.07%
IQAEvelNet	80.60%

Table 11 – Classification results for each of the image enhancement algorithms on the NIH Chest X-ray dataset. Bold numbers highlight the best results.

Algorithm	Accuracy
None	**73.52%**
UM	71.42%
HEF	72.55%
CLAHE	72.47%
ATACE	70.85%
TCDHE	72.55%
IQAEvelNet	70.36%

impact on image classification tasks. This is likely because the labels used for training were created based on the perception of medical professionals analyzing non-enhanced images. Therefore, the experiment does not account for potential changes in visual inspection diagnostic accuracy that could result from improved image quality. To more effectively demonstrate the impact of enhancement algorithms, it would be better to compare visual inspection diagnostic accuracy between unimproved and improved image datasets. This approach would directly measure how these algorithms influence the interpretation of medical professionals. The diagnostic accuracy using each dataset could then be compared to more accurate diagnostic tools, such as biopsies or blood tests, to assess the true effectiveness of the enhanced images in improving visual inspection diagnostic outcomes.

Thus, the proposed method emerges as a superior alternative to state-of-the-art algorithms in terms of NIQE score, while maintaining reasonable entropy and computational cost. It can be used as a effective pre-processing tool for chest X-ray image enhancement to support visual inspection diagnostics.

Table 12 – Evolutionary results for weight set 1. ↑ means that higher values are better, ↓ means the opposite.

Epoch	NIQE ↓	Time (seconds) ↓
1	3.18	1329.21
2	3.18	1329.11
3	3.18	1329.59
4	3.18	1326.22

Table 13 – Evolutionary results for weight set 2. ↑ means that higher values are better, ↓ means the opposite.

Epoch	Entropy ↑	Time (seconds) ↓
1	7.937	239.25
2	7.937	239.29
3	7.949	239.45
4	7.948	240.13
5	7.948	239.82
6	7.948	239.64
7	7.948	239.79

4.6 Evolutionary process

In this section we will discuss the changes of the parameters during the evolutionary process and the computational cost. The IQA metrics and NIQE values presented here are referent to the validation dataset and the processing time is for the entire epoch training time. Table 12 shows the NIQE and processing time for weight set 1, where NIQE weight = 1 and the others are 0. For this weight set, the optimal solution was found on the first epoch and after 3 epoch without improvement the algorithm stopped.

Table 13 shows the entropy and processing time for weight set 2, where entropy weight = 1 and the others are 0. For this weight set, the optimal solution was found on the fourth epoch. We can see that from the third to the fourth epoch there was a drop in entropy from 7.949 to 7.948 and the algorithm considered it a improvement. Because the algorithm judges using the training dataset an apparent improvement can, when dealing with small margins, generate a deterioration in the validation results. Comparing the processing time results with the previous weight set we can see a considerable difference. The quickest epoch of weight set 1 was 1326.22 seconds while the quickest epoch of weight set 2 was 239.25. As the only difference between both processes is the fitness calculation process, we can infer that most on the computational cost is spent there.

Table 14 shows the NIQE, entropy and processing time for weight set 3, where NIQE

Table 14 – Evolutionary results for weight set 3. ↑ means that higher values are better, ↓ means the opposite.

Epoch	NIQE ↓	Entropy ↑	Time (seconds) ↓
1	3.75	7.761	1420.75
2	3.75	7.761	1410.50
3	3.75	7.761	1408.58
4	3.75	7.761	1410.24

weight = 1, entropy weight = 1 and SSIM weight = 0. For this weight set, the optimal solution was found on the first epoch. Comparing the processing time results with weight set 1 we can see a small increase. The quickest epoch of weight set 3 was 1408.58 seconds while the quickest epoch of weight set 1 was 1326.22. This increase can be attributed to the fact that weight set 3 uses both NIQE and entropy for fitness calculation while weight set 1 uses only NIQE.

Table 15 shows the NIQE, entropy, SSIM and processing time for weight set 4, where NIQE weight = 1, entropy weight = 1 and SSIM weight = 1. For this weight set, the optimal solution was found on the tenth epoch. Comparing the processing time results with weight set 3 we can see a small increase. The quickest epoch of weight set 4 was 1621.16 seconds while the quickest epoch of weight set 3 was 1408.58. This increase can be attributed to the fact that weight set 4 includes SSIM in it's fitness calculation. The solution on the first epoch had SSIM of 0.680, NIQE of 3.85 and entropy of 7.530. In the next epoch improved the SSIM score, from 0.680 to 0.855, and deteriorated the NIQE and entropy score, from 3.86 to 4.02 and from 7.530 to 7.370 respectively. The next evolution happened on epoch 5. In this epoch the SSIM score was reduced but the NIQE and entropy score improved, from 4.02 to 3.75 and from 7.370 to 7.423 respectively. The last evolution occured in epoch 7. In this case, all scores improved. SSIM went from 0.787 to 0.800, NIQE from 3.75 to 3.58, and entropy from 7.423 to 7.457.

Table 16 shows the NIQE, entropy, SSIM and processing time for weight set 5, where NIQE weight = 1, entropy weight = 1 and SSIM weight = 2. For this weight set, the optimal solution was found on the tenth epoch. Comparing the processing time results with weight set 4 we can see they are similar. Because both weight set use the three scores, their execution process is similar. The weights only influence the final multiplication. The final result has a worst NIQE, from 3.58 to 3.90, and entropy, from 7.457 to 7.241, when compared to weight set 4. To compensate this it has a better SSIM, from 0.800 to 0.850. We can observe that the higher weight for SSIM incentives the algorithm to choose higher SSIM solutions. This preference comes at the cost of NIQE and entropy scores.

Table 15 – Evolutionary results for weight set 4. ↑ means that higher values are better, ↓ means the opposite.

Epoch	SSIM ↑	NIQE ↓	Entropy ↑	Time (seconds) ↓
1	0.680	3.86	7.530	1627.50
2	0.855	4.02	7.370	1641.45
3	0.855	4.02	7.370	1621.16
4	0.855	4.02	7.370	1639.59
5	0.787	3.75	7.423	1636.11
6	0.787	3.75	7.423	1625.63
7	0.800	3.58	7.457	1631.91
8	0.800	3.58	7.457	1646.29
9	0.800	3.58	7.457	1641.46
10	0.800	3.58	7.457	1646.90

Table 16 – Evolutionary results for weight set 5. ↑ means that higher values are better, ↓ means the opposite.

Epoch	SSIM ↑	NIQE ↓	Entropy ↑	Time (seconds) ↓
1	0.850	4.43	7.462	1635.01
2	0.884	4.41	7.396	1623.34
3	0.884	4.41	7.396	1642.70
4	0.855	4.06	7.458	1627.24
5	0.855	4.06	7.458	1644.91
6	0.855	4.06	7.458	1570.55
7	0.850	3.90	7.241	1563.70
8	0.850	3.90	7.241	1546.45
9	0.850	3.90	7.241	1544.84
10	0.850	3.90	7.241	1582.59

Table 17 – Evolutionary results for weight set 6. ↑ means that higher values are better, ↓ means the opposite.

Epoch	SSIM ↑	NIQE ↓	Entropy ↑	Time (seconds) ↓
1	0.738	3.86	7.492	1564.37
2	0.804	3.96	7.535	1552.46
3	0.804	3.96	7.535	1547.15
4	0.804	3.96	7.535	1540.50
5	0.804	3.96	7.535	1555.99

Table 17 shows the NIQE, entropy, SSIM and processing time for weight set 6, where NIQE weight = 1, entropy weight = 2 and SSIM weight = 1. For this weight set, the optimal solution was found on the tenth epoch. The final result has a worst NIQE, from 3.58 to 3.96, a similar SSIM, from 0.800 to 0.804, and a better entropy, from 7.457 to 7.535 when compared to weight set 4. We can observe that the higher weight for entropy incentives the algorithm to choose higher entropy solutions. This preference comes at the cost of the NIQE score.

Table 18 – Evolutionary results for weight set 7. ↑ means that higher values are better, ↓ means the opposite.

Epoch	SSIM ↑	NIQE ↓	Entropy ↑	Time (seconds) ↓
1	0.713	3.47	7.489	1549.28
2	0.713	3.47	7.489	1530.34
3	0.713	3.47	7.489	1532.00
4	0.713	3.47	7.489	1537.80

Table 19 – Evolutionary results for weight set 8. ↑ means that higher values are better, ↓ means the opposite.

Epoch	SSIM ↑	NIQE ↓	Entropy ↑	Time (seconds) ↓
1	0.669	4.35	7.634	1542.23
2	0.843	4.10	7.530	1578.74
3	0.843	4.10	7.530	1558.14
4	0.843	4.10	7.530	1543.05
5	0.843	4.10	7.530	1545.44

Table 18 shows the NIQE, entropy, SSIM and processing time for weight set 7, where NIQE weight = 2, entropy weight = 1 and SSIM weight = 1. For this weight set, the optimal solution was found on the first epoch. The final result has a better NIQE, from 3.58 to 3.47, a better entropy, from 7.457 to 7.489, and a worst SSIM, from 0.800 to 0.713, when compared to weight set 4. We can observe that the higher weight for NIQE incentives the algorithm to choose higher NIQE solutions. This preference comes at the cost of the SSIM score.

Table 19 shows the NIQE, entropy, SSIM and processing time for weight set 8, where NIQE weight = 1, entropy weight = 4 and SSIM weight = 1. It was expected that by increasing the entropy weight from 2 to 4, in relation to weight set 6, the result should have a better entropy score and lose in another score to compensate. For this weight set, the optimal solution was found on the second epoch. The final result has a worst NIQE, from 3.96 to 4.10, a slightly worst entropy, from 7.535 to 7.530, and a better SSIM, from 0.804 to 0.843, when compared to weight set 6. The experimental result showed an increase in SSIM and a slight decrease in entropy. This shows that while the weights influence the final result it is not deterministic. In this case, the algorithm found a solution with a slightly smaller entropy score but with a significantly bigger SSIM score.

Table 20 shows the NIQE, entropy, SSIM and processing time for weight set 9, where NIQE weight = 2, entropy weight = 4 and SSIM weight = 1. It was expected that by increasing the SSIM weight from 1 to 2, in relation to weight set 8, the result should have a better SSIM score and lose in another score to compensate. For this weight set, the optimal

Table 20 – Evolutionary results for weight set 9. ↑ means that higher values are better, ↓ means the opposite.

Epoch	SSIM ↑	NIQE ↓	Entropy ↑	Time (seconds) ↓
1	0.854	3.97	7.608	1535.93
2	0.854	3.97	7.608	1525.14
3	0.854	3.97	7.608	1530.34
4	0.854	3.97	7.608	1522.06

Table 21 – Evolutionary results for weight set 10. ↑ means that higher values are better, ↓ means the opposite.

Epoch	SSIM ↑	NIQE ↓	Entropy ↑	Time (seconds) ↓
1	0.685	4.20	7.720	1545.23
2	0.648	4.01	7.761	1531.37
3	0.648	4.01	7.761	1531.60
4	0.648	4.01	7.761	1527.19
5	0.648	4.01	7.761	1541.21

solution was found on the first epoch. The final result has a better NIQE, from 4.10 to 3.97, a better entropy, from 7.530 to 7.608, and a better SSIM, from 0.843 to 0.854, when compared to weight set 8. The experimental result showed an improvement in all scores. This reinforces the non-deterministic impact of weights in the final result. In this case, the algorithm found a solution in all parameters.

Table 21 shows the NIQE, entropy, SSIM and processing time for weight set 10, where NIQE weight = 1, entropy weight = 8 and SSIM weight = 1. It was expected that by reducing the SSIM weight from 2 to 1, in relation to weight set 9, and increasing the entropy weight from 4 to 8 the result should have a better entropy score and a worst SSIM score. For this weight set, the optimal solution was found on the second epoch. The final result has a worst NIQE, from 3.97 to 4.01, a better entropy, from 7.608 to 7.761, and a worst SSIM, from 0.854 to 0.648, when compared to weight set 9. The experimental result showed the expected outcome.

Table 22 shows the NIQE, entropy, SSIM and processing time for weight set 11, where NIQE weight = 1, entropy weight = 16 and SSIM weight = 1. It was expected that by increasing the entropy weight from 8 to 16, in relation to weight set 10, the result should have a better entropy score and a worst NIQE and/or entropy score. For this weight set, the optimal solution was found on the third epoch. The final result has a better NIQE, from 4.01 to 3.61, a better entropy, from 7.761 to 7.798, and a worst SSIM, from 0.648 to 0.643, when compared to weight set 10. The experimental result showed the expected outcome.

Table 22 – Evolutionary results for weight set 11. ↑ means that higher values are better, ↓ means the opposite.

Epoch	SSIM ↑	NIQE ↓	Entropy ↑	Time (seconds) ↓
1	0.688	4.22	7.773	1513.38
2	0.688	4.22	7.773	1515.87
3	0.643	3.61	7.798	1518.99
4	0.643	3.61	7.798	1519.00
5	0.643	3.61	7.798	1515.45
6	0.643	3.61	7.798	1523.60

5 CONCLUSION

X-ray imaging is commonly used in medicine due to its cost-effectiveness and non-invasive nature, but it frequently encounters noise problems stemming from equipment or environmental influences. While there are state-of-the-art algorithms that aim to enhance the quality of X-ray images, most of them use non-learning approaches. To address this issue this work proposes an unsupervised evolutionary algorithm for X-ray image enhancement: IQAEvolNet. This evolutionary algorithm trains a base convolutional neural network model, LightdehazeTestnet, through selection, permutation and mutation. The proposed method uses the no-reference Image Quality Assessment algorithms Natural Image Quality Evaluator (NIQE) and Entropy, together with Structural Similarity Index Measure, for fitness calculation. IQAEvolNet achieved a better NIQE score of 4.05 compared to 4.24 and a faster processing time of 2.95 milliseconds compared to 0.195 seconds, in comparison to the state-of-the-art algorithm with the best NIQE and entropy, HEF. The proposed method outperformed state-of-the-art algorithms in NIQE score and showed competitiveness in entropy score and processing time.

At last, for future works, it is suggest the research of image quality assessment metrics specialized in X-ray images. In our experiments it is demonstrated that better NIQE and entropy results do not imply a higher quality image. Furthermore, a more in-depth research can be done to develop a CNN more suited to chest X-ray image enhancement.

5.1 Acknowledgement

This work was supported by CNPq with grants No. 313599/2019-0 and No. 305517/2022-8. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

REFERENCES

AHN, E.; KUMAR, A.; FULHAM, M.; FENG, D.; KIM, J. Unsupervised domain adaptation to classify medical images using zero-bias convolutional auto-encoders and context-based feature augmentation. **IEEE Transactions on Medical Imaging**, v. 39, n. 7, p. 2385–2394, 2020.

AURANGZEB, K.; ASLAM, S.; ALHUSSEIN, M.; NAQVI, R. A.; ARSALAN, M.; HAIDER, S. I. Contrast enhancement of fundus images by employing modified pso for improving the performance of deep learning models. **IEEE Access**, v. 9, p. 47930–47945, 2021.

BRUCE, M.; HANNAH, A.; HAMMOND, R.; KHAING, Z. Z.; TREMBLAY-DARVEAU, C.; BURNS, P. N.; HOFSTETTER, C. P. High-frequency nonlinear doppler contrast-enhanced ultrasound imaging of blood flow. **IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control**, v. 67, n. 9, p. 1776–1784, 2020.

CHOKCHAITHANAKUL, W.; PUNYABUKKANA, P.; CHUANGSUWANICH, E. Adaptive image preprocessing and augmentation for tuberculosis screening on out-of-domain chest x-ray dataset. **IEEE Access**, v. 10, p. 132144–132152, 2022.

CHOWDHURY, M. E. H.; RAHMAN, T.; KHANDAKAR, A.; MAZHAR, R.; KADIR, M. A.; MAHBUB, Z. B.; ISLAM, K. R.; KHAN, M. S.; IQBAL, A.; EMADI, N. A.; REAZ, M. B. I.; ISLAM, M. T. Can ai help in screening viral and covid-19 pneumonia? **IEEE Access**, v. 8, p. 132665–132676, 2020.

DING, B.; ZHANG, R.; XU, L.; LIU, G.; YANG, S.; LIU, Y.; ZHANG, Q. U2d2net: Unsupervised unified image dehazing and denoising network for single hazy image enhancement. **IEEE Transactions on Multimedia**, v. 26, p. 202–217, 2024.

DING, W.; CHAKRABORTY, S.; MALI, K.; CHATTERJEE, S.; NAYAK, J.; DAS, A. K.; BANERJEE, S. An unsupervised fuzzy clustering approach for early screening of covid-19 from radiological images. **IEEE Transactions on Fuzzy Systems**, v. 30, n. 8, p. 2902–2914, 2022.

FAN, G.-D.; FAN, B.; GAN, M.; CHEN, G.-Y.; CHEN, C. L. P. Multiscale low-light image enhancement network with illumination constraint. **IEEE Transactions on Circuits and Systems for Video Technology**, v. 32, n. 11, p. 7403–7417, 2022.

GAZDA, M.; PLAVKA, J.; GAZDA, J.; DROTáR, P. Self-supervised deep convolutional neural network for chest x-ray classification. **IEEE Access**, v. 9, p. 151972–151982, 2021.

GUO, W.; WANG, H.; HAN, Z.; ZHONG, J.; LIU, Z. A novel low-light catenary image enhancement approach for cscs detection in high-speed railways. **IEEE Open Journal of Instrumentation and Measurement**, v. 1, p. 1–7, 2022.

HADIKHANI, P.; LAI, D. T. C.; ONG, W.-H. Human activity discovery with automatic multi-objective particle swarm optimization clustering with gaussian mutation and game theory. **IEEE Transactions on Multimedia**, v. 26, p. 420–435, 2024.

HAYATI, M.; MUCHTAR, K.; ROSLIDAR; MAULINA, N.; SYAMSUDDIN, I.; ELWIREHARDJA, G. N.; PARDAMEAN, B. Impact of clahe-based image enhancement for diabetic retinopathy classification through deep learning. **Procedia Computer Science**, v. 216, p. 57–66, 2023. ISSN 1877-0509. 7th International Conference on Computer Science and Computational Intelligence 2022. Available at: <https://www.sciencedirect.com/science/article/pii/S1877050922021895>.

HOWARD, A. G.; ZHU, M.; CHEN, B.; KALENICHENKO, D.; WANG, W.; WEYAND, T.; ANDREETTO, M.; ADAM, H. **MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications**. 2017.

HU, J.; GUO, X.; CHEN, J.; LIANG, G.; DENG, F.; LAM, T. L. A two-stage unsupervised approach for low light image enhancement. **IEEE Robotics and Automation Letters**, v. 6, n. 4, p. 8363–8370, 2021.

HUANG, D.; LIU, J.; ZHOU, S.; TANG, W. Deep unsupervised endoscopic image enhancement based on multi-image fusion. **Computer Methods and Programs in Biomedicine**, v. 221, p. 106800, 2022. ISSN 0169-2607. Available at: <https://www.sciencedirect.com/science/article/pii/S0169260722001742>.

JAEGER, S.; CANDEMIR, S.; ANTANI, S.; WáNG, Y.-X. J.; LU, P.-X.; THOMA, G. Two public chest x-ray datasets for computer-aided screening of pulmonary diseases. **Quantitative Imaging in Medicine and Surgery**, v. 4, n. 6, 2014. ISSN 2223-4306. Available at: <https://qims.amegroups.org/article/view/5132>.

JALALI, S. M. J.; AHMADIAN, S.; KHOSRAVI, A.; SHAFIE-KHAH, M.; NAHAVANDI, S.; CATALãO, J. P. S. A novel evolutionary-based deep convolutional neural network model for intelligent load forecasting. **IEEE Transactions on Industrial Informatics**, v. 17, n. 12, p. 8243–8253, 2021.

JAMIL, S.; PIRAN, M. J.; KWON, O.-J. A comprehensive survey of transformers for computer vision. **Drones**, v. 7, n. 5, 2023. ISSN 2504-446X. Available at: <https://www.mdpi.com/2504-446X/7/5/287>.

JIANG, Q.; MAO, Y.; CONG, R.; REN, W.; HUANG, C.; SHAO, F. Unsupervised decomposition and correction network for low-light image enhancement. **IEEE Transactions on Intelligent Transportation Systems**, v. 23, n. 10, p. 19440–19455, 2022.

JIANG, Y.; GONG, X.; LIU, D.; CHENG, Y.; FANG, C.; SHEN, X.; YANG, J.; ZHOU, P.; WANG, Z. Enlightengan: Deep light enhancement without paired supervision. **IEEE Transactions on Image Processing**, v. 30, p. 2340–2349, 2021.

KANDULA, P.; SUIN, M.; RAJAGOPALAN, A. N. Illumination-adaptive unpaired low-light enhancement. **IEEE Transactions on Circuits and Systems for Video Technology**, v. 33, n. 8, p. 3726–3736, 2023.

KUMAR, M.; BHANDARI, A. K. Novel unsupervised learning architecture for exposure-based classification and enhancement. **IEEE Transactions on Artificial Intelligence**, v. 4, n. 5, p. 1064–1075, 2023.

KUMAR, S.; BHANDARI, A. K. Automatic tissue attenuation-based contrast enhancement of low-dynamic x-ray images. **IEEE Transactions on Radiation and Plasma Medical Sciences**, v. 6, n. 5, p. 574–582, 2022.

KUMAR, S.; BHANDARI, A. K.; RAJ, A.; SWARAJ, K. Triple clipped histogram-based medical image enhancement using spatial frequency. **IEEE Transactions on NanoBioscience**, v. 20, n. 3, p. 278–286, 2021.

LEPCHA, D. C.; GOYAL, B.; DOGRA, A.; SHARMA, K. P.; GUPTA, D. N. A deep journey into image enhancement: A survey of current and emerging trends. **Information Fusion**, v. 93, p. 36–76, 2023. ISSN 1566-2535. Available at: <https://www.sciencedirect.com/science/article/pii/S1566253522002603>.

LI, D.; BIAN, Z.; LI, S.; HE, J.; ZENG, D.; MA, J. Noise characteristics modeled unsupervised network for robust ct image reconstruction. **IEEE Transactions on Medical Imaging**, v. 41, n. 12, p. 3849–3861, 2022.

LI, X.; HE, R.; WU, J.; YAN, H.; CHEN, X. Lees-net: Fast, lightweight unsupervised curve estimation network for low-light image enhancement and exposure suppression. **Displays**, v. 80, p. 102550, 2023. ISSN 0141-9382. Available at: <https://www.sciencedirect.com/science/article/pii/S014193822300183X>.

LIANG, J.; XU, Y.; QUAN, Y.; SHI, B.; JI, H. Self-supervised low-light image enhancement using discrepant untrained network priors. **IEEE Transactions on Circuits and Systems for Video Technology**, v. 32, n. 11, p. 7332–7345, 2022.

LIN, C.-H.; WU, J.-X.; LI, C.-M.; CHEN, P.-Y.; PAI, N.-S.; KUO, Y.-C. Enhancement of chest x-ray images to improve screening accuracy rate using iterated function system and multilayer fractional-order machine learning classifier. **IEEE Photonics Journal**, v. 12, n. 4, p. 1–18, 2020.

LIN, F.; ZHANG, H.; WANG, J.; WANG, J. Unsupervised image enhancement under non-uniform illumination based on paired cnns. **Neural Networks**, v. 170, p. 202–214, 2024. ISSN 0893-6080. Available at: <https://www.sciencedirect.com/science/article/pii/S089360802300638X>.

LIN, Q.; FANG, Z.; CHEN, Y.; TAN, K. C.; LI, Y. Evolutionary architectural search for generative adversarial networks. **IEEE Transactions on Emerging Topics in Computational Intelligence**, v. 6, n. 4, p. 783–794, 2022.

LIU, K.; TIAN, Y. Research and analysis of deep learning image enhancement algorithm based on fractional differential. **Chaos, Solitons & Fractals**, v. 131, p. 109507, 2020. ISSN 0960-0779. Available at: <https://www.sciencedirect.com/science/article/pii/S096007791930459X>.

LU, Z.; WHALEN, I.; DHEBAR, Y.; DEB, K.; GOODMAN, E. D.; BANZHAF, W.; BODDETI, V. N. Multiobjective evolutionary design of deep convolutional neural networks for image classification. **IEEE Transactions on Evolutionary Computation**, v. 25, n. 2, p. 277–291, 2021.

LUO, Y.; MAJOE, S.; KUI, J.; QI, H.; PUSHPARAJAH, K.; RHODE, K. Ultra-dense denoising network: Application to cardiac catheter-based x-ray procedures. **IEEE Transactions on Biomedical Engineering**, v. 68, n. 9, p. 2626–2636, 2021.

LUO, Y.; YOU, B.; YUE, G.; LING, J. Pseudo-supervised low-light image enhancement with mutual learning. **IEEE Transactions on Circuits and Systems for Video Technology**, v. 34, n. 1, p. 85–96, 2024.

MA, Y.; LIU, J.; LIU, Y.; FU, H.; HU, Y.; CHENG, J.; QI, H.; WU, Y.; ZHANG, J.; ZHAO, Y. Structure and illumination constrained gan for medical image enhancement. **IEEE Transactions on Medical Imaging**, v. 40, n. 12, p. 3955–3967, 2021.

MALALI, H. E.; ASSIR, A.; BHATEJA, V.; MOUHSEN, A.; HARMOUCHI, M. A contrast enhancement model for x-ray mammograms using modified local s-curve transformation based on multi-objective optimization. **IEEE Sensors Journal**, v. 21, n. 10, p. 11543–11554, 2021.

MITTAL, A.; SOUNDARARAJAN, R.; BOVIK, A. C. Making a "completely blind" image quality analyzer. **IEEE Signal Processing Letters**, v. 20, n. 3, p. 209–212, 2013.

MUNADI, K.; MUCHTAR, K.; MAULINA, N.; PRADHAN, B. Image enhancement for tuberculosis detection using deep learning. **IEEE Access**, v. 8, p. 217897–217907, 2020.

PRAJAPATI, K.; CHUDASAMA, V.; PATEL, H.; UPLA, K.; RAJA, K.; RAMACHANDRA, R.; BUSCH, C. Direct unsupervised super-resolution using generative adversarial network (dus-gan) for real-world data. **IEEE Transactions on Image Processing**, v. 30, p. 8251–8264, 2021.

QI, Y.; GUO, Y.; WANG, Y. Image quality enhancement using a deep neural network for plane wave medical ultrasound imaging. **IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control**, v. 68, n. 4, p. 926–934, 2021.

RAHMAN, T.; KHANDAKAR, A.; QIBLAWEY, Y.; TAHIR, A.; KIRANYAZ, S.; Abul Kashem, S. B.; ISLAM, M. T.; Al Maadeed, S.; ZUGHAIER, S. M.; KHAN, M. S.; CHOWDHURY, M. E. Exploring the effect of image enhancement techniques on covid-19 detection using chest x-ray images. **Computers in Biology and Medicine**, v. 132, p. 104319, 2021. ISSN 0010-4825. Available at: <https://www.sciencedirect.com/science/article/pii/S001048252100113X>.

RAHMAN, T.; KHANDAKAR, A.; QIBLAWEY, Y.; TAHIR, A.; KIRANYAZ, S.; Abul Kashem, S. B.; ISLAM, M. T.; Al Maadeed, S.; ZUGHAIER, S. M.; KHAN, M. S.; CHOWDHURY, M. E. Exploring the effect of image enhancement techniques on covid-19 detection using chest x-ray images. **Computers in Biology and Medicine**, v. 132, p. 104319, 2021. ISSN 0010-4825. Available at: <https://www.sciencedirect.com/science/article/pii/S001048252100113X>.

SHAMSHAD, F.; KHAN, S.; ZAMIR, S. W.; KHAN, M. H.; HAYAT, M.; KHAN, F. S.; FU, H. Transformers in medical imaging: A survey. **Medical Image Analysis**, v. 88, p. 102802, 2023. ISSN 1361-8415. Available at: <https://www.sciencedirect.com/science/article/pii/S1361841523000634>.

SHAO, D.; LU, X.; LIU, X. 3d intracranial aneurysm classification and segmentation via unsupervised dual-branch learning. **IEEE Journal of Biomedical and Health Informatics**, v. 27, n. 4, p. 1770–1779, 2023.

SHI, Y.; WANG, B.; WU, X.; ZHU, M. Unsupervised low-light image enhancement by extracting structural similarity and color consistency. **IEEE Signal Processing Letters**, v. 29, p. 997–1001, 2022.

SUN, L.; WANG, J.; HUANG, Y.; DING, X.; GREENSPAN, H.; PAISLEY, J. An adversarial learning approach to medical image synthesis for lesion detection. **IEEE Journal of Biomedical and Health Informatics**, v. 24, n. 8, p. 2303–2314, 2020.

TIAN, Y.; LIU, R.; ZHANG, X.; MA, H.; TAN, K. C.; JIN, Y. A multipopulation evolutionary algorithm for solving large-scale multimodal multiobjective optimization problems. **IEEE Transactions on Evolutionary Computation**, v. 25, n. 3, p. 405–418, 2021.

ULLAH, H.; MUHAMMAD, K.; IRFAN, M.; ANWAR, S.; SAJJAD, M.; IMRAN, A. S.; ALBUQUERQUE, V. H. C. de. Light-dehazenet: A novel lightweight cnn architecture for single image dehazing. **IEEE Transactions on Image Processing**, v. 30, p. 8968–8982, 2021.

WANG, L.; LIU, Y.; MI, J.; ZHANG, J. Mse-fusion: Weakly supervised medical image fusion with modal synthesis and enhancement. **Engineering Applications of Artificial Intelligence**, v. 119, p. 105744, 2023. ISSN 0952-1976. Available at: <https://www.sciencedirect.com/science/article/pii/S0952197622007345>.

WANG, R.; JIANG, B.; YANG, C.; LI, Q.; ZHANG, B. Magan: Unsupervised low-light image enhancement guided by mixed-attention. **Big Data Mining and Analytics**, v. 5, n. 2, p. 110–119, 2022.

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: **2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**. IEEE, 2017. Available at: <http://dx.doi.org/10.1109/CVPR.2017.369>.

WANG, Y.; GUO, J.; GAO, H.; YUE, H. Uiec²-net: Cnn-based underwater image enhancement using two color space. **Signal Processing: Image Communication**, v. 96, p. 116250, 2021. ISSN 0923-5965. Available at: <https://www.sciencedirect.com/science/article/pii/S0923596521001004>.

WANG, Z.; BOVIK, A.; SHEIKH, H.; SIMONCELLI, E. Image quality assessment: from error visibility to structural similarity. **IEEE Transactions on Image Processing**, v. 13, n. 4, p. 600–612, 2004.

WANG, Z.; LI, C.; MO, Y.; SHANG, S. Rca-cyclegan: Unsupervised underwater image enhancement using red channel attention optimized cyclegan. **Displays**, v. 76, p. 102359, 2023. ISSN 0141-9382. Available at: <https://www.sciencedirect.com/science/article/pii/S0141938222001779>.

WEN, Y.-W.; PENG, S.-H.; TING, C.-K. Two-stage evolutionary neural architecture search for transfer learning. **IEEE Transactions on Evolutionary Computation**, v. 25, n. 5, p. 928–940, 2021.

WU, Z.; ZHOU, Q.; WANG, F. Coarse-to-fine lung nodule segmentation in ct images with image enhancement and dual-branch network. **IEEE Access**, v. 9, p. 7255–7262, 2021.

WU, Z.; ZHU, F.; GUO, K.; SHENG, R.; CHAO, L.; FANG, H. Modal adaptive super-resolution for medical images via continual learning. **Signal Processing**, v. 217, p. 109342, 2024. ISSN 0165-1684. Available at: <https://www.sciencedirect.com/science/article/pii/S0165168423004164>.

XIE, X.; ZHANG, X.; YE, S.; XIONG, D.; OUYANG, L.; YANG, B.; ZHOU, H.; WAN, Y. Mrscfusion: Joint residual swin transformer and multiscale cnn for unsupervised multimodal medical image fusion. **IEEE Transactions on Instrumentation and Measurement**, v. 72, p. 1–17, 2023.

XU, B.; ZHOU, D.; LI, W. Image enhancement algorithm based on gan neural network. **IEEE Access**, v. 10, p. 36766–36777, 2022.

YADAV, P.; MENON, N.; RAVI, V.; VISHVANATHAN, S. Lung-gans: Unsupervised representation learning for lung disease classification using chest ct and x-ray images. **IEEE Transactions on Engineering Management**, v. 70, n. 8, p. 2774–2786, 2023.

YAN, S.; CHEN, X.; WU, Z.; TAN, M.; YU, J. Hybrur: A hybrid physical-neural solution for unsupervised underwater image restoration. **IEEE Transactions on Image Processing**, v. 32, p. 5004–5016, 2023.

YANG, H.; HUYAN, J.; MA, T.; TONG, Z.; HAN, C.; XIE, T. Novel computer tomography image enhancement deep neural networks for asphalt mixtures. **Construction and Building Materials**, v. 352, p. 129067, 2022. ISSN 0950-0618. Available at: <https://www.sciencedirect.com/science/article/pii/S0950061822027222>.

YANG, S.; TIAN, Y.; HE, C.; ZHANG, X.; TAN, K. C.; JIN, Y. A gradient-guided evolutionary approach to training deep neural networks. **IEEE Transactions on Neural Networks and Learning Systems**, v. 33, n. 9, p. 4861–4875, 2022.

YE, S.; LI, Z.; MCCANN, M. T.; LONG, Y.; RAVISHANKAR, S. Unified supervised-unsupervised (super) learning for x-ray ct image reconstruction. **IEEE Transactions on Medical Imaging**, v. 40, n. 11, p. 2986–3001, 2021.

YU, Y.-F.; ZHONG, G.; ZHOU, Y.; CHEN, L. Fs-gan: Fuzzy self-guided structure retention generative adversarial network for medical image enhancement. **Information Sciences**, v. 642, p. 119114, 2023. ISSN 0020-0255. Available at: <https://www.sciencedirect.com/science/article/pii/S0020025523006990>.

ZHANG, M.; LI, H.; PAN, S.; LYU, J.; LING, S.; SU, S. Convolutional neural networks-based lung nodule classification: A surrogate-assisted evolutionary algorithm for hyperparameter optimization. **IEEE Transactions on Evolutionary Computation**, v. 25, n. 5, p. 869–882, 2021.

ZHANG, R.; YAO, S.; LU, L.; WANG, X. Scnet: A self-calibrating unsupervised low-light image enhancement network. **IEEE Sensors Journal**, v. 23, n. 24, p. 30765–30772, 2023.

ZHANG, Y.; JIANG, Q.; LIU, P.; GAO, S.; PAN, X.; ZHANG, C. Underwater image enhancement using deep transfer learning based on a color restoration model. **IEEE Journal of Oceanic Engineering**, v. 48, n. 2, p. 489–514, 2023.

ZHOU, W.; LIU, Y.; LI, M.; WANG, Y.; SHEN, Z.; FENG, L.; ZHU, Z. Dynamic multi-objective optimization framework with interactive evolution for sequential recommendation. **IEEE Transactions on Emerging Topics in Computational Intelligence**, v. 7, n. 4, p. 1228–1241, 2023.

ZHOU, X.; ZHANG, J.; ZHOU, F. Underwater image enhancement method based on dynamic heterogeneous feature fusion neural network. **IEEE Access**, v. 10, p. 91816–91827, 2022.

ZHOU, Y.; YEN, G. G.; YI, Z. A knee-guided evolutionary algorithm for compressing deep neural networks. **IEEE Transactions on Cybernetics**, v. 51, n. 3, p. 1626–1638, 2021.

ZHU, Z.; WEI, H.; HU, G.; LI, Y.; QI, G.; MAZUR, N. A novel fast single image dehazing algorithm based on artificial multiexposure image fusion. **IEEE Transactions on Instrumentation and Measurement**, v. 70, p. 1–23, 2021.