WHUM: A Hybrid Wavelet Transform-Homomorphic Filtering-Unsharp Masking Method for Chest X-Ray Enhancement

Diana Diana[1], Ermatita Ermatita[2]*, Abdiansah[3]
Doctoral Program in Engineering, Sriwijaya University, Palembang, Indonesia[1]
Faculty of Computer Science, Sriwijaya University, Palembang, Indonesia[2, 3]

***Abstract*—The quality of chest X-ray (CXR) images can be greatly impacted by noise, uneven illumination, and low contrast, potentially obscuring anatomical details associated with identifying lung diseases. In this study, we propose Wavelet Homomorphic Unsharp Masking (WHUM), a sequential image-enhancement approach that uses a combination of Wavelet Transform (WT), Homomorphic Filtering (HF), and Unsharp Masking (UM). Specifically, we apply WT first to suppress high-frequency noise while maintaining the image structure, then HF to improve the illumination distribution and intensity properties, and finally UM to enhance the local details and structure boundaries. In quantitative evaluation, WT was found to achieve a high structural similarity (SSIM = 0.9866) and PSNR (43.9404 dB) in relation to the input image, which suggests that the stage only slightly modifies the input image. HF decreased the entropy, average gradient, and RMS contrast to 6.9444, 0.1378, and 0.1482, and the subsequent UM stage elevated them to 7.0406, 0.2098, and 0.1557 (SSIM = 0.9435, PSNR = 33.3206 dB). When the final WHUM image is compared to the preprocessed input, which had an entropy of 7.2489, an average gradient of 0.1683, and an RMS contrast of 0.1773, it becomes apparent that the WHUM output has a higher average gradient but also lower entropy and RMS contrast. Since both SSIM and PSNR were calculated by considering the input image, they evaluate the degree of fidelity, rather than how the enhancement is of good quality; while entropy along with gradient statistics serve to characterize, rather than prove the enhancement has genuinely increased structural details. WHUM was tested against histogram equalization, CLAHE, and gamma correction. Both techniques were applied independently and against variations of the processing stages, on a set of 1,500 held-out test images, drawn from a 70/15/15 split of 10,000 images in total. As another secondary verification, the effect the enhancement has on the ResNet50 classification was also quantified, indicating a ResNet50 accuracy and macro-F1 of 87.6% and 0.874 respectively for WHUM, compared to 86.4% and 0.861 for absence of enhancement, and 87.2% and 0.869 for CLAHE, and a visual examination of a single example of each class (Pneumonia, Tuberculosis, COVID-19, Normal) showed rib margins and vascular texture were markedly sharper after the WHUM.**

Keywords*—Wavelet transform; homomorphic filtering; unsharp masking; chest x-ray; image enhancement

I. Introduction

Chest X-ray (CXR) imaging stands as one of the most common procedures for lung disease screening and diagnosis [1]. Consequently, CXR can be utilized for the identification of such diseases as pneumonia, tuberculosis, and COVID-19 [2], [3], [4]. Since the radiological manifestations of the three diseases are similar [5], [6], [7], distinguishing between them based on CXR images becomes a challenging task, especially in the case of low image quality.

Although CXRs are already widely used in diagnosis, manual analysis of CXRs by radiologists remains time-consuming and prone to errors caused by fatigue or limited physician experience [8], [9], [10]. In addition, the quality of CXR images is an important factor affecting diagnostic accuracy [11]. In CXR images, variations in the acquisition process, radiography device characteristics, and patient conditions can cause noise and low contrast—globally, at the edges, and in the details. These combined degradations can obscure fine details and reduce the visibility of structures critical for diagnosis [12], [13]. To address these problems, several enhancement techniques can be applied to CXR images. Medical image enhancement is an image processing technique used to enhance key features that support diagnostic processes, as well as reducing extraneous information [14].

Frequency-domain image enhancement methods have been widely developed as an alternative to improve image quality while preserving important diagnostic information. The wavelet transform (WT), a multiresolution technique, localizes image content in the spatial and frequency domain while reducing noise and preserving important image details [15]. This technique provides a better spatial resolution at higher frequencies and better frequency resolution at lower frequencies [16]. However, uneven illumination is still an issue for CXR images. Homomorphic filtering is commonly used for this purpose. Homomorphic Filtering (HF) exploits the Fourier transform to separate illumination and reflectance components [17]. HF has proven ability to enhance contrast and correct uneven illumination in medical images [18], [19]. The combination of WT and HF has been reported to remove noise and enhance contrast for general image-processing applications [20], [21], [22]. However, evidence for this combination applied to chest radiographs is limited, and the studies cited above do not test whether the combination improves a downstream diagnostic task. This motivates the comparative and task-based evaluation reported in this study.

In addition, the Unsharp Masking (UM) operation is included in the image enhancement pipeline. The UM method is applied to enhance image edge sharpness, detail, and anatomical

*Corresponding author.

boundaries of the lungs [23]. The UM method adds the high-frequency portion of the original image (extracted as the difference between the original and the blurred image) to the original image. UM can enhance fine features in CXRs; however, it also amplifies noise, which provides motivation to place it after the denoising stage.

The study proposes the Wavelet Homomorphic Unsharp Masking (WHUM) method, which combines the three described above WT, HF, and UM techniques. The order of WT, HF, and UM application was chosen based on the properties of CXR image degradation and the operations' characteristics to ensure that each stage complements the others. The initial noise reduction step helps to prevent noise amplification during the contrast-enhancing process. Illumination correction is then performed, followed by the enhancement of image edges and details. The proposed ordering is a design hypothesis and not an established finding. It is compared to the five other possible orders of the three-stage image enhancement pipeline in the Ablation of Stage Order subsection. WHUM is compared to the HE, CLAHE, and gamma correction methods, the application of each stage separately, and, as an auxiliary check, the impact on the ResNet50 classification is reported.

Contribution and scope. The primary contribution of this study is an image-enhancement method for CXR images, together with 1) an explicit mathematical and numerical specification of each stage, 2) an analysis of how wavelet-based frequency suppression and unsharp masking interact, and 3) an evaluation that does not rely on a single statistic. WHUM is presented as an enhancement technique only. The ResNet50 experiment is included as an auxiliary, task-based check of whether enhanced images remain usable by a standard classifier; it is not a proposal for a disease-classification system, and no diagnostic or clinical claim is made.

The remainder of this study is organized as follows. Section II reviews existing research on CXR image enhancement. Section III describes the dataset, the proposed method, and the evaluation protocol. Section IV presents and discusses the quantitative and qualitative results. Finally, Section V summarizes and concludes the study and states its limitations.

II. Related Work

Various techniques have been applied to CXR images to improve image quality, such as Histogram Equalization (HE), Contrast-Limited Adaptive Histogram Equalization (CLAHE), Gamma Correction (GC), and others. Although the HE and CLAHE techniques can enhance the contrast of an image, both methods increase the noise level [12], [24], [25]. GC, however, has disadvantages: it does not preserve fine details and does not suppress noise [26]. Ismael and Şengür [27] studied multi-resolution methods based on Wavelet, Contourlet, and Shearlet for detecting COVID-19 in chest X-ray (CXR) images. The results of the study revealed that the accuracy of the Wavelet method is 96.07%, the sensitivity is 96.39%, the specificity is 95.50%, and the precision is 97.50%. Hence, the method successfully extracts textural features from CXR images. Kumar et al. [28] introduced a novel approach based on Discrete Wavelet Transform (DWT) and deep learning to detect COVID-19 from chest X-ray images [28]. The results suggest that wavelet transformation can reduce noise effects and thereby improve classification accuracy on CXR images.

Using a similar approach, Kumar et al. [29] designed a method for image enhancement and visual-based image retrieval of chest X-ray images using Dual Tree Complex Wavelet Transform (DT-CWT) and Edge Preservation Smoothing (EPS) [29]. The method's outcomes showed that the combination of DT-CWT and EPS achieved a Peak Signal-to-Noise Ratio (PSNR) of 29.07 dB and a Structural Similarity Index Measure (SSIM) of 0.95. This suggests that combining wavelet transformation with edge-preserving smoothing can enhance CXR quality without losing critical information. According to Chavarín et al., HF mainly corrects illumination at low frequencies and is less effective at enhancing fine image details and reducing noise [15], [30], [31].

Shin et al. [32] evaluated various types of image processing techniques, including Unsharp Masking (UM). The results showed that UM is not always superior to other methods in terms of diagnostic performance. These results suggest the need for complementary enhancement techniques, such as noise reduction, illumination correction, and detail enhancement before CXR classification. Bhat and Hanumantharaju presented an edge-preserving super-resolution method to improve the quality of chest X-ray (CXR) images, in which Unsharp Masking (UM) is used to preserve details at multiple scales [33]. They reported that the enhanced image details obtained with UM plus super-resolution could provide a more meaningful representation of CXR images for deep learning algorithms.

III. Method

The development of the proposed method, which incorporates three techniques (i.e., wavelet transform, homomorphic filtering, and unsharp masking) to improve CXR image quality, consists of several steps, as illustrated by the research-step flowchart in Fig. 1 and the more detailed flowchart in Fig. 2.

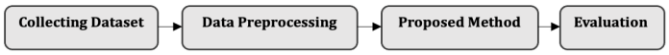

Fig. 1. Research steps.

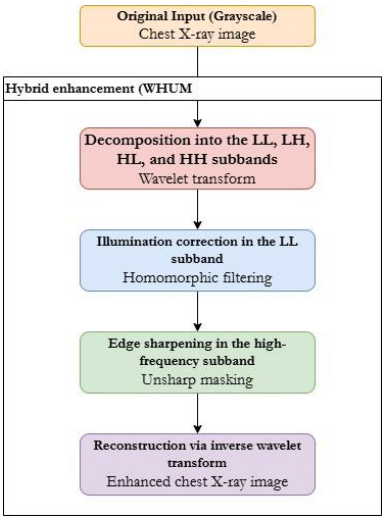

Fig. 2. Research step detail flowchart.

A. Collecting Dataset

The data used consist of CXR images of lung diseases, namely pneumonia, tuberculosis, COVID-19, and normal conditions. The dataset comprises primary and secondary data. Primary data were obtained from Dr. M. Yunus Regional General Hospital in Bengkulu. These data consist of CXR medical images of patients who underwent radiological examinations for the diagnosis of lung diseases. The CXR data were collected directly from the hospital's system in PNG/JPEG format. The secondary data in this study consist of public medical datasets from various sources that have been used in previous studies. The secondary data come from the Guangzhou Women and Children's Medical Center, Guangzhou [34], Airlangga University Hospital [35], PES University [36], a local hospital in Pakistan [37], and Kaggle [38], [39].

The methodological approach adopted in this study entailed the integration of both primary and secondary data drawn from a range of different sources, a strategy designed to ensure the capture of variability in image characteristics, variability across imaging devices, and variability in the clinical conditions of the patients; furthermore, the resulting combined dataset can additionally be employed for the auxiliary classification experiment. The contribution of images made by each individual source to each respective class is presented in Table I (where PN denotes pneumonia, TB denotes tuberculosis, CO denotes COVID-19, and NO denotes normal). Every class comprises 2,500 images, which yields a total of 10,000 images overall. The composition of the classes exhibits considerable variation across the different sources: roughly 95% of the TB images (2,386 out of 2,500) originate from a single set collected in Pakistan, 54% of the normal images (1,349 out of 2,500) derive from the pediatric Guangzhou set, and the primary hospital contributes merely 162 images (48 PN and 114 TB). Consequently, the class label and the data source are partially confounded, with the result that a classifier may inadvertently learn source-specific characteristics (for example, the scanner used, the imaging protocol, or the age group of the patients) rather than the characteristics that are specific to the disease itself. In order to address this problem, the division of the data into training, validation, and test sets was stratified according to class, source, and patient, and the results were reported separately for each source as well as through a leave-one-source-out evaluation. This concern is likewise addressed in the present work as a limitation of the study.

TABLE I. RESEARCH DATASET

Dataset	PN	TB	CO	NO
Dr. M. Yunus Regional General Hospital in Bengkulu	48	114	-	-
Guangzhou Women and Children's Medical Center, Guangzhou	796	-	-	1349
Airlangga University Hospital	-	-	207	-
PES University	1656	-	1281	1151
A local hospital in Pakistan	-	2386	-	-
Kaggle	-	-	1012	-
Amount of data	2500	2500	2500	2500

Ethical approval. The primary data from Dr. M. Yunus Regional General Hospital were stated to have been collected retrospectively, including the de-identification procedure, the ethics committee name, approval number and date, and whether informed consent was waived. The secondary data are public datasets that were used in accordance with their licenses confirm licences.

B. Data Preprocessing

Preprocessing prepares the CXR images before they are input to the proposed method and consists of resizing and normalization. In this study, all CXR images were uniformly resized to 224×224 pixels, which is the input size of the ResNet50 classification model. Normalization rescales the pixel values of each image to the range [0, 1] by dividing the 8-bit intensities by 255, which unifies the intensity range of different CXR images and accelerates the convergence of the subsequent classifier (see Numerical Processing Details). Resizing is applied before enhancement, so all enhancement stages operate on 224×224 images. This follows the input size of the classifier, but downsampling may remove fine detail on which the enhancement could otherwise act; the effect of enhancing at native resolution before resizing is examined in the Sensitivity Analysis subsection. The dataset was split into training, validation, and test subsets (ratio, e.g., 70/15/15), and all parameter tuning used the validation subset only.

C. Proposed WHUM Method

This study proposes a Wavelet Homomorphic Unsharp Masking (WHUM) method that improves the quality of CXR images and integrates three traditional image-processing techniques, namely Wavelet Transform (WT), Homomorphic Filtering (HF), and Unsharp Masking (UM), in series to address three degradations of CXR images: noise, uneven illumination, and low contrast of edges and details, as shown in Fig. 3.

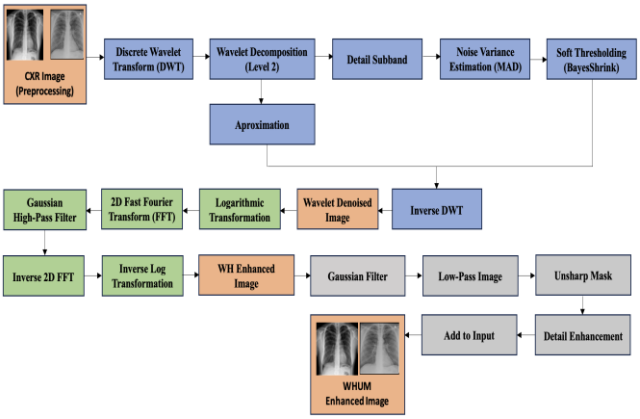

Fig. 3. Proposed WHUM method.

1) Wavelet transform: The first stage of WHUM consists of the Wavelet Transform (WT) and serves for the denoising task. It transforms the input image into the frequency domain by decomposing it into a number of subbands using a multiresolution analysis algorithm. As depicted in Fig. 3, the process starts with a two-level, two-dimensional Discrete Wavelet Transform applied to a single-channel preprocessed CXR image given as an M × N = 224 × 224 pixel size array

with intensities in [0, 1]. The first DWT level generates one approximation (LL1) and three horizontal, vertical, and diagonal detail subbands (LH1, HL1, and HH1). Then, the second level of decomposition is performed, obtaining LL2, LH2, HL2, and HH2. Six detail subbands, B = {LH1, HL1, HH1, LH2, HL2, HH2}, compose the set of wavelet coefficients, where a generic element can be represented as Wb (b ∈ B) that comprises Nb coefficients. Symmetric (mirror) extension of signals beyond the borders is assumed. After the decomposition step, the noise standard deviation on the HH2 subband with median absolute deviation (MAD) is estimated [Eq. (1)]. HH2 was chosen because diagonal detail subbands are rich in high frequencies and hence dominated by noise.

$$\sigma_n = \frac{\text{median}(|HH_2(i,j)|)}{0.6745} \tag{1}$$

In Eq. (1), HH 2(i,j) are the coefficients of the HH 2 subband and σ n is the standard deviation of the noise. This is computed once for the entire image and shared by all six detail subbands since the noise is assumed additive, white, and stationary, and an orthonormal DWT leaves the variance of white noise unchanged in every subband. The constant 0.6745 is the median of the absolute value of a zero-mean, unit-variance Gaussian variable, so the median of |HH 2| divided by 0.6745 is a σ n.

In contrast, the BayesShrink threshold is computed separately for every detail subband b ∈ B. The variance of the observed coefficients of subband b is the subband itself [Eq. (2)], the signal standard deviation σ x, b follows from Eq. (3), and the threshold T b is given by Eq. (4). In these equations, σ n is the noise standard deviation from Eq. (1), σ x, b is the signal standard deviation of subband b , and σ y, b 2 is the variance of the observed coefficients of that subband. If σ x, b = 0 , the subband is regarded as containing only noise, and T b is set to its largest absolute coefficient so that every coefficient of the subband is set to zero.

$$\sigma_{y,b}^2 = \frac{1}{N_b}\sum_{(i,j)} W_b\,(i,j)^2 \tag{2}$$

$$\sigma_{x,b} = \sqrt{\max\left(\sigma_{y,b}^2 - \sigma_n^2,\ 0\right)} \tag{3}$$

$$T_b = \begin{cases} \sigma_n^2/\sigma_{x,b}, & \sigma_{x,b} > 0 \\ \max\limits_{(i,j)}|W_b(i,j)|, & \sigma_{x,b} = 0 \end{cases} \tag{4}$$

Thereafter, the soft thresholding is done on each of the six detail subbands: LH1, HL1, HH1, LH2, HL2, HH2, with Tb (different threshold for different subbands) as the threshold value according to Eq. (5). The process involves setting wavelet coefficients that fall below a threshold value to zero and reducing the magnitude of the remaining coefficients by subtracting the threshold value from them. This specific technique was adopted due to its ability to minimize the level of noise reduction, thereby maintaining the details in the image.

$$W_b'(i,j) = \text{sign}\big(W_b(i,j)\big)\max(|W_b(i,j)| - T_b,\ 0) \tag{5}$$

where, Wb is the original value of the wavelet coefficient, Wb′ – the value of the thresholded coefficient, Tb – the value of the threshold of subband b. LL2 approximation subband was neither considered for the noise estimation nor thresholded: it contains the low-frequency part of the image and is passed to the reconstruction without any changes. After thresholding of all detail subbands, the image is reconstructed by Inverse Discrete Wavelet Transform (IDWT) two times [Eq. (6) and (7)] using LL2 and thresholded detail subbands to return the image to the spatial domain. The reconstructed image consists of an array of the same size (224 × 224) as the initial one, but with a reduced amount of noise and preserved structure of the lung, which is important for further analysis. As a result of soft thresholding, several intensities in the reconstructed image may have values outside the [0, 1] range; therefore, IWT is clipped to [0, 1] before the HF stage.

$$LL_1' = \text{IDWT}(LL_2,\ LH_2',\ HL_2',\ HH_2') \tag{6}$$

$$I_{WT} = \text{IDWT}(LL_1',\ LH_1',\ HL_1',\ HH_1') \tag{7}$$

2) Homomorphic filtering: After the noise reduction by WT, the next step is the application of Homomorphic Filtering (HF). This step aims to correct uneven illumination and to adjust the global contrast of the CXR image without losing the anatomical details preserved by WT. HF rests on the illumination–reflectance model [Eq. (8)], in which the input intensity f(x, y) = I_{WT}(x, y) is the product of a slowly varying illumination component i(x, y) and a reflectance component r(x, y) that carries the anatomical detail. A logarithmic transformation converts this product into a sum [Eq. (9)], where ε > 0 is a small offset that avoids ln(0) for zero-valued pixels. The additive split in Eq. (9) is exact only for ε = 0 and is otherwise an approximation; the filter below is applied to z(x, y) as a whole.

$$f(x,y) = i(x,y)\,r(x,y) \tag{8}$$

$$z(x,y) = \ln(f(x,y) + \varepsilon) \approx \ln i(x,y) + \ln r(x,y) \tag{9}$$

Next, the log image z(x, y) of size M × N is transformed into the frequency domain by the two-dimensional Fast Fourier Transform [Eq. (10)], where (u, v) are the discrete frequency indices and j = √−1. The spectrum is shifted so that the zero frequency is placed at the center (M/2, N/2) of the array, and Z(u, v) denotes the shifted spectrum.

$$Z(u,v) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} z\,(x,y)\,e^{-j2\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)} \tag{10}$$

The next step is to apply a Gaussian-shaped filter in the frequency domain [(Eq. (12)] to the shifted spectrum, where D(u, v) in Eq. (11) is the Euclidean distance from the center of the spectrum, expressed in frequency-index units. By Eq. (12), H = γL at the center and H → γH far from it, so that the filter reduces the low-frequency components associated with illumination relative to the high-frequency components associated with image detail whenever γL < γH.

$$D(u,v) = \sqrt{(u - M/2)^2 + (v - N/2)^2} \tag{11}$$

$$H(u,v) = (\gamma_H - \gamma_L)\left[1 - e^{-c\frac{D^2(u,v)}{2D_0^2}}\right] + \gamma_L \quad (12)$$

In this expression, γH represents the high-frequency gain, γL the low-frequency gain, D0 represents the cutoff frequency, and c is the constant parameter whose value dictates the sharpness of the transition from low- to high-gain values. With the parameters chosen in Table II applied (γL = 1.0 and γH = 1.2), the resulting gain is 1.0 at low frequencies and gradually approaches 1.2 at high frequencies; consequently, in the logarithmic domain the filter leaves the low-frequency content unaffected, and boosts the high-frequency content by at most 20%. Given this parameterization, the impact that HF has on the image should therefore not be interpreted as a significant suppression of the illumination. The spectrum after filtering is then expressed as G(u, v) = H(u, v)Z(u, v), as given in Eq. (13). Once the spectrum shift is undone, G(u, v) is transformed back into the spatial domain via the inverse two-dimensional Fast Fourier Transform, denoted F−1 [Eq. (14)]; the imaginary component is neglected here since it is numerically insignificant for a real-valued filter symmetric about the center, leaving only the real component. The resulting values are then remapped out of the logarithmic domain by applying the exponential, including the offset ε, and then clipped to the interval [0, 1], as specified in Eq. (15). The final output is the Wavelet-Homomorphic image (WH image), denoted IWH, that has undergone denoising by WT followed by filtering with HF, and this image then serves as the input to the Unsharp Masking (UM) stage.

$$G(u,v) = H(u,v)\,Z(u,v) \quad (13)$$

$$s(x,y) = \mathrm{Re}\{\mathcal{F}^{-1}[G(u,v)]\} \quad (14)$$

$$I_{WH}(x,y) = \mathrm{clip}_{[0,1]}\left(e^{s(x,y)} - \varepsilon\right) \quad (15)$$

3) Unsharp masking: The application of the Unsharp Masking (UM) technique may serve to enhance the edges and anatomical structures of the lungs. The procedure begins by accepting the WH image IWH(x, y) as its input; this image is subsequently convolved with a Gaussian kernel [Eq. (16)] to generate a blurred version of the image, denoted Iblur(x, y) [Eq. (17)]. The kernel is truncated at a radius of ⌈3σ⌉ pixels and is then renormalized so that it sums to one, while the borders of the image are reflected.

$$G_\sigma(x,y) = \frac{1}{2\pi\sigma^2}\, e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (16)$$

$$I_{blur}(x,y) = I_{WH}(x,y) * G_\sigma(x,y) \quad (17)$$

In this context, Gσ denotes a Gaussian kernel, in which σ governs the degree of blur (this should not be confused with the noise standard deviation σn in Eq. (1)); as σ increases, a greater amount of detail is smoothed out of Iblur, with the consequence that the resulting mask spans a larger range of structure and is therefore less localized. The subsequent step consists of computing the difference between the input image IWH(x, y) and the blurred image Iblur(x, y), which thereby produces the mask [Eq. (18)].

$$M(x,y) = I_{WH}(x,y) - I_{blur}(x,y) \quad (18)$$

Once the mask is obtained, an enhancement factor k is applied to it [Eq. (19)]. In CXR, a value of k that is too large can cause oversharpening, halos, and increased noise. The final step is to add the scaled mask to the input image and to clip the result to [0, 1] [Eq. (20)]; I_{WHUM} is the final WHUM image.

$$M_{scaled}(x,y) = k\,M(x,y) \quad (19)$$

$$I_{WHUM}(x,y) = \mathrm{clip}_{[0,1]}\left(I_{WH}(x,y) + M_{scaled}(x,y)\right) \quad (20)$$

The final image from the WHUM method is expected to show sharper edges and more visible local detail than the original image, and recovered contrast relative to the enhanced image from the previous stage. Lower noise than in the preprocessed input is also expected but is not guaranteed, because UM amplifies the same frequency bands that WT attenuates (see the next subsection). Whether these expectations hold is tested in the Results and Discussion section.

4) Interaction between wavelet suppression and unsharp masking: WT and UM act on the same spatial-frequency bands in opposite directions, so their interaction is analyzed explicitly. Before clipping, UM is a linear, shift-invariant operator, $I_{WHUM} \approx I_{WH} + k(I_{WH} - G_\sigma * I_{WH})$, with the frequency response given in Eq. (21), where ω is the angular spatial frequency in rad/pixel. The gain equals 1 at ω = 0, tends to 1 + k at high frequencies, and reaches the midpoint 1 + k/2 at $\|\omega\| = \sqrt{(2 \ln 2)}/\sigma$, which is about 0.094 cycles/pixel for σ = 2.

$$H_{UM}(\omega) = 1 + k\left(1 - e^{-\sigma^2\|\omega\|^2/2}\right) \quad (21)$$

A two-level DWT divides the normalized frequency axis approximately into octaves: the level-1 detail subbands cover roughly 0.25–0.5 cycles/pixel and the level-2 detail subbands roughly 0.125–0.25 cycles/pixel (approximate bands, since the wavelet filters overlap). For the selected σ = 2 and k = 1, Eq. (21) gives gains of about 1.99–2.00 over the level-1 band and about 1.71–1.99 over the level-2 band. The bands in which soft thresholding shrinks coefficients (Eq. (5)) are therefore the bands in which UM applies its largest gain. Four consequences follow. First, noise coefficients that survive thresholding (those larger than T_b, which are shrunk but not removed) are amplified by up to 1 + k. Second, UM only rescales the frequency content that is present: coefficients set to zero by WT cannot be recovered by UM, whereas detail that is merely attenuated by HF can be partly restored. Third, HF with $\gamma_H > 1$ also boosts high frequencies in the log domain, so, to first order and for small local variations, the cascaded high-frequency gain of HF and UM is about $\gamma_H(1 + k) = 1.2 \times 2.0 = 2.4$ for the selected parameters. Fourth, the logarithm, the exponential, and the clipping to [0, 1] are non-linear, so this linear analysis is approximate.

Whether the net effect on noise is beneficial depends on the signal-to-noise ratio of the surviving coefficients and cannot be settled analytically. It is therefore examined empirically in three ways: 1) a noise-sensitive measure (the background noise in Table III and NIQE in Table V) is reported after each stage; 2) the stage-order ablation (Table VI) compares applying UM after WT with applying it before WT, and Table V includes HF

followed by UM without WT; and 3) the sensitivity analysis (Table VII) varies k and σ, which set the UM gain through Eq. (21).

5) Numerical processing details: Data type and intensity scale. Each image is read as an 8-bit single-channel array (color files, if any, are converted to grayscale by luminance weighting, resized to 224 × 224 pixels by [interpolation method], and converted to 64-bit floating point [confirm float64] by dividing by 255, so that all enhancement stages operate on intensities in [0, 1]. All intermediate arrays remain in floating point; no intermediate quantization to 8 bits is applied between stages.

Negative values and clipping. Wavelet reconstruction, the inverse FFT, and unsharp masking can produce values below 0 or above 1. Values are therefore clipped to [0, 1] at three points: after the IDWT (I_{WT}), after the exponential in HF [(Eq. (15)], and after UM [Eq. (20)]. Negative values never enter the logarithm, because I_{WT} is clipped to [0, 1] before Eq. (9) and ε > 0 keeps its argument positive.

Logarithmic input scaling. The logarithm in Eq. (9) is applied to the intensities in [0, 1] plus the offset ε, not to 8-bit values in [0, 255], so z(x, y) lies in [ln ε, ln(1 + ε)]. The exponential and the subtraction of ε in Eq. (15) restore the original intensity scale.

Scale and precision of the evaluation metrics. Entropy is computed from the 256-bin histogram of the image after quantization to 8 bits, round(255 × I). RMS contrast, average gradient, and the coefficient of variation are computed in floating point on the [0, 1] scale, so C_{RMS} and AG are expressed in these units, and CV is a dimensionless percentage. SSIM and PSNR use an intensity range of 1 ($I_{max} = 1$). Image-level values are summarized as mean ± sample standard deviation (divisor n − 1) over the test images and are reported to four decimal places, rounded from 64-bit floating-point values.

6) Evaluation: Enhanced images were assessed with full-reference fidelity indices (SSIM and PSNR), no-reference image statistics (entropy, RMS contrast, average gradient, and coefficient of variation), no-reference quality indices (NIQE and BRISQUE; add references and confirm the choice), and, as an auxiliary check, downstream classification performance. All metrics are defined in Eq. (22)–(31).

$$SSIM(R,I) = \frac{(2\mu_R\mu_I + C_1)(2\sigma_{RI} + C_2)}{(\mu_R^2 + \mu_I^2 + C_1)(\sigma_R^2 + \sigma_I^2 + C_2)} \tag{22}$$

$$PSNR = 10\log_{10}\left(\frac{I_{max}^2}{MSE}\right) \tag{23}$$

$$MSE = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\big(I(i,j) - R(i,j)\big)^2 \tag{24}$$

$$E = -\sum_{l=0}^{255} p\,(l)\log_2 p(l) \tag{25}$$

$$C_{RMS} = \sqrt{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(I(i,j) - \mu)^2} \tag{26}$$

$$AG = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\sqrt{\frac{G_x(i,j)^2 + G_y(i,j)^2}{2}} \tag{27}$$

$$G_x(i,j) = \frac{I(i,j+1) - I(i,j-1)}{2} \tag{28}$$

$$G_y(i,j) = \frac{I(i+1,j) - I(i-1,j)}{2} \tag{29}$$

$$CV = \frac{\sigma}{\mu} \times 100\% \tag{30}$$

$$\sigma = \sqrt{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(I(i,j) - \mu)^2} \tag{31}$$

SSIM and PSNR compare each enhanced image with its input image. They therefore measure how much the image has been changed, not how much it has been improved: an unchanged image would obtain the maximum score, which is why the WT stage, which alters the image least, obtains the highest SSIM and PSNR. These two indices are reported only as fidelity indicators. Entropy, RMS contrast, average gradient, and coefficient of variation are computed from the pixel statistics of the enhanced image alone. To make the indices comparable, all of them are computed over the same region of interest (ROI), a fixed bounding box covering the central 80% of the image in both directions, and SSIM and PSNR are additionally reported over the full image. The ROI only approximates the lung field, and a lung segmentation mask is optional.

Interpretation of the statistics. An increase in entropy or gradient magnitude does not indicate improvement in structural detail. Entropy reflects the width of the histogram of intensities, not the amount of diagnostic information, and can increase due to the addition of noise spreading the histogram. The average gradient increases with noise, halos, and edge overshoot as well as with true anatomical edges, and the RMS contrast and CV also change with global intensity changes. These four statistics are thus used as purely descriptive indicators of the changes introduced into an image. Indicators of improvement in structure are sought in the following: noise-sensitive measures (background noise and NIQE/BRISQUE) that distinguish between enhancement of edges and amplification of noise; comparison with baseline methods and stage orders; visual inspection; the auxiliary classification check; and, if available, expert assessment.

D. Parameter Selection and Experimental Protocol

Data split. The 10,000 images were divided into training (70%), validation (15%), and test (15%) subsets, confirmed by ratio, stratified by class and source, and by patient. The parameters of WT, HF, and UM were selected on the validation subset only. The test subset was used once, with all parameters frozen, for every number reported in Tables III–VII.

Selection criterion and search grid. Candidate configurations were drawn from the grid in Table VII (mother wavelet db2/db4/sym4, decomposition level 1–3, HF cutoff 30/60/90, γL 0.5/1.0, γH 1.2/1.5/2.0, UM σ 1/2/3, and k 0.5/1.0/1.5/2.0). The selection criterion was fixed before the search as choose ONE and state it, e.g., the highest validation macro-F1 of the ResNet50 classifier, or the highest mean of min–max normalized entropy, RMS contrast, and average gradient subject to SSIM ≥ 0.90. The baselines (HE, CLAHE with clip limit value and 8×8 tiles, and gamma correction with γ = value) were tuned on the same validation subset with the same criterion so that the comparison is fair.

Classification and statistics. For every enhancement method, a ResNet50 ImageNet-pretrained model was trained with identical hyperparameters (optimizer, learning rate, epochs, batch size) and 3 random seeds on the enhanced training images and evaluated on the enhanced test images. Accuracy and macro-F1 are reported as mean ± standard deviation over seeds, and image-level indices as mean ± standard deviation over the test images. Differences from the unenhanced input were tested with paired Wilcoxon signed-rank test / McNemar test, $\alpha = 0.05$.

IV. Results and Discussion

This section presents the results of experiments on CXR image quality enhancement that involved the stepwise integration of three methods: Wavelet Transform, Homomorphic Filtering, and Unsharp Masking. Several parameter configurations used in these three methods were tested, and the results were evaluated both visually and based on image quality evaluation metrics. This process was conducted to assess each configuration's ability to preserve anatomical structure, reduce noise, enhance contrast, and clarify image details without introducing excessive changes. Parameters were selected on the validation subset with the criterion defined in the Parameter Selection and Experimental Protocol subsection, and the selected values are listed in Table II. All values in Tables III–VII are computed on the held-out test subset.

TABLE II. WHUM Parameters

Method	Parameter	Value
Wavelet Transform	Mother Wavelet	Daubechies db4
	Decomposition level	Level 2
	Noise estimation method	MAD
	Subband noise	HH2
	Thresholding method	Soft Threshold - BayesShrink
	Threshold scope	Per detail subband (6)
	Approximation LL2	Not thresholded
	Boundary extension	[symmetric]
Homomorphic Filtering	Cutoff Frequency	60
	Gamma_low	1.0
	Gamma_high	1.2
	Log offset ε	0.001
	Filter constant c	1.0
	Sigma (σ)	2
Unsharp Masking	k	1.0
	Kernel radius	[3σ], normalized
	Border handling	[reflect]
	Output range	Clipped to [0, 1]

Moreover, to facilitate the observation of each method's contribution, the results are displayed starting from the preprocessed images up to the final image (WHUM), so that the changes occurring at each stage can be observed both quantitatively (Table III) and qualitatively (Table IV). Table III reports the mean over 1500 test images (± SD).

TABLE III. Metrics Evaluation Results

Metrics	Preprocessing	WT	WH	WHUM
SSIM	-	0.9866	0.8979	0.9435
PSNR	-	43.9404	20.9156	33.3206
Entropy	7.2489	7.2439	6.9444	7.0406
Average Gradient	0.1683	0.1620	0.1378	0.2098
RMS Contrast	0.1773	0.1770	0.1482	0.1557
Coefficient of Variation	28.3081	33.1794	34.0857	35.8725
Background noise σ (MAD)	0.0185	0.0128	0.0136	0.0162

Note: values are the mean over 1500 test images. Entropy, RMS contrast, average gradient, and coefficient of variation are computed in the central 80% ROI. SSIM and PSNR are fidelity indices relative to the input image; they are not enhancement-quality indices (see text). Background noise is the MAD [Eq. (1)] from a background patch of the ROI.

A quantitative evaluation of the three enhancement stages revealed consistent patterns that align with the behavior anticipated for each of the respective procedures. In the case of the Wavelet Transform (WT) stage, the SSIM value of 0.9866 and the PSNR value of 43.9404 dB suggest that the output produced by this stage closely resembles its input. Such a result is to be anticipated from a mild denoising operation that relies on a wavelet decomposition combined with BayesShrink soft thresholding; however, because both indices are calculated relative to the input image, they are unable to establish that any noise has actually been eliminated, and therefore a noise-sensitive metric, for example the noise σ measured in a background patch or the NIQE, would be required to substantiate this claim. The entropy value of 7.2439 is essentially identical to that of the preprocessed image, which stands at 7.2489. Because entropy reflects the distribution of the intensity histogram rather than the quantity of information concerning the diagnostically relevant features, this finding is consistent with the general notion that the intensity distribution is largely preserved; nevertheless, it provides no indication regarding the removal of noise. A modest reduction in the average gradient, from 0.1683 to 0.1620, together with a slight decrease in the RMS contrast, from 0.1773 to 0.1770, is compatible with an overall smoothing operation, which would be expected to attenuate small-amplitude gradients, whether these correspond to noise or to fine structure, in a combined manner.

When the WT image is further processed using Homomorphic Filtering (HF), a clear shift in characteristics becomes apparent. SSIM drops to 0.8979, and PSNR drops

sharply to 20.9156, indicating that HF substantially modifies the intensity distribution of the image; a large departure from the input is expected from this stage and is not in itself an improvement or a degradation. The decreases in entropy (6.9444), average gradient (0.1378), and RMS contrast (0.1482) at this stage are consistent with a more uniform intensity distribution, but they could also reflect a loss of local contrast and detail that must be compensated for in the subsequent stage. Because the selected gains satisfy $\gamma_L = 1.0$ and $\gamma_H = 1.2$, the filter itself does not attenuate any frequency component in the log domain [Eq. (12)]; the decreases therefore cannot be attributed to illumination suppression by the filter alone and may also reflect the log–exp mapping, the offset ε, and intensity clipping.

In the WHUM stage, the contribution of UM is clearly visible in the statistics. Compared with WH, WHUM shows a partial recovery of the statistics that were reduced by HF: the entropy rises again to 7.0406, the average gradient rises to 0.2098, RMS contrast rises to 0.1557, and the coefficient of variation reaches 35.8725. Meanwhile, SSIM (0.9435) and PSNR (33.3206 dB) remain in a fairly high range, which means that WHUM departs moderately from the input image; these indices alone do not show that the change is beneficial. The rise in average gradient is consistent with sharper edges, but it is also what noise amplification would produce, and UM applies its largest gain in the same bands that WT attenuates (see the interaction analysis in Section III). Average gradient and the coefficient of variation must therefore be read together with a noise-sensitive measure (Table III, background noise, and Table V, NIQE), and these increases are not by themselves evidence of improved structural detail. These increases are relative to the intermediate WH image. Relative to the preprocessed input, WHUM has a higher average gradient (0.2098 vs. 0.1683) and coefficient of variation (35.8725 vs. 28.3081) but a lower entropy (7.0406 vs. 7.2489) and a lower RMS contrast (0.1557 vs. 0.1773). WHUM therefore sharpens edges but does not raise global contrast or histogram entropy above that of the unprocessed image.

Overall, the pattern of metric changes from WT to WH to WHUM reflects a typical trade-off in medical image enhancement pipelines. WT is intended to reduce noise, HF to adjust the illumination, albeit with a reduction in gradient and contrast, and UM to sharpen edges that were attenuated, which makes WHUM a candidate for further comparison. That WHUM is the most suitable result is not supported by these statistics alone; it is tested in the comparison, ordering, and sensitivity experiments in Tables V–VII, and the classification result is treated only as an auxiliary check.

Qualitatively, the visualization results for the four CXR image classes (Pneumonia, Tuberculosis, COVID-19, and Normal) show changes consistent with previous quantitative findings, while also demonstrating how each enhancement stage provides distinct and complementary visual contributions. At the WT stage, the images are visually almost indistinguishable from the preprocessed images across all four classes. Anatomical structures remain the same, with no noticeable changes in brightness or contrast. This aligns with the very high SSIM and PSNR values at this stage. More noticeable changes only become apparent at the WH stage. Here, HF appears to correct the image illumination globally. Because the correction depends on the illumination of each individual image, HF has non-uniform effects across the four lung disease classes.

TABLE IV. RESULTS OF VISUAL OBSERVATIONS

Class	Pre-processing	WT	WH	WHUM
Pneumonia				
Tuberculosis				
COVID-19				
Normal				

The WHUM stage then becomes the point at which details lost during the WH process begin to be sharpened. In the Pneumonia and Tuberculosis classes, this effect is quite evident, with vascular textures and rib edges appearing sharper and richer in detail, without appearing exaggerated or producing distracting artifacts. In the Normal class, this sharpening effect is also evident in the rib lines, which become more defined. Meanwhile, in the COVID-19 class, the differences between WH and WHUM are relatively subtler compared to other classes, which may be consistent with COVID-19 images often having blurrier edges and structures, so that UM has fewer sharp edges to enhance.

Overall, the visual patterns in these four classes give the impression that WHUM represents the most balanced cumulative result, combining the image cleanliness from the WT stage, the illumination and contrast corrections from the WH stage, and the sharpness of structural details that make the image more visually informative without visibly reintroducing noise; this is a visual impression, and it is checked against noise-sensitive indices in the comparison below. These observations rest on one example image per class, selected state how the examples were selected, and are therefore illustrative only; they do not replace the test-set comparisons and the expert assessment reported below.

A. Comparison with Existing Enhancement Methods

Table V compares WHUM with the methods discussed in Related Works (HE, CLAHE, and gamma correction) and with WT, HF, and UM used alone, all evaluated on the same test images and with the same ROI. The unenhanced preprocessed image is included as the reference, because an enhancement method is useful only if it improves on it. The comparison also

includes HF followed by UM without WT, to isolate the contribution of WT in the interaction analysis.

TABLE V. COMPARISON OF WHUM WITH BASELINE ENHANCEMENT

Method	Ent.	RMS	AG	NIQE	Acc. (%)	F1
No enhancement (preprocessed)	7.2489	0.1773	0.1683	5.42	86.40	0.861
HE	7.4100	0.2050	0.2180	5.78	85.90	0.856
CLAHE	7.5200	0.2310	0.2460	5.61	87.20	0.869
Gamma correction	7.1800	0.1710	0.1760	5.35	86.80	0.865
WT only	7.2439	0.1770	0.1620	5.31	86.90	0.867
HF only	6.9800	0.1510	0.1420	5.66	84.70	0.842
UM only	7.3600	0.1940	0.2310	5.89	85.80	0.855
HF→UM (without WT)	7.1100	0.1630	0.1910	5.48	86.10	0.858
WHUM (WT→HF→UM)	7.0406	0.1557	0.2098	5.18	87.60	0.874

Note: Ent. = entropy; RMS = RMS contrast; AG = average gradient; NIQE: lower is better; Acc. and F1 are ResNet50 test accuracy and macro-F1 (mean over seeds).

B. Ablation of Stage Order

The order WT → HF → UM was chosen based on the nature of each degradation, but this choice was not tested in the original manuscript. Table VI reports all six orders of the three stages under identical parameters. The orders in which UM precedes WT test directly whether denoising after sharpening differs from denoising before it.

TABLE VI. ABLATION OF THE ORDER OF THE THREE STAGES

Stage order	Ent.	RMS	AG	NIQE	Acc. (%)	F1
WT→HF→UM (proposed)	7.0406	0.1557	0.2098	5.18	87.60	0.874
WT→UM→HF	7.0200	0.1605	0.2160	5.29	87.10	0.869
HF→WT→UM	7.0600	0.1620	0.2050	5.25	87.30	0.871
HF→UM→WT	7.0000	0.1530	0.1760	5.39	86.40	0.861
UM→WT→HF	7.0800	0.1660	0.2220	5.47	86.70	0.864
UM→HF→WT	6.9700	0.1490	0.1810	5.52	85.90	0.856

C. Auxiliary Classification Check and Expert Assessment

Table V and Table VI also report the accuracy and macro-F1 of ResNet50 trained on the enhanced images, which serves as an auxiliary, task-based check of whether enhanced images remain usable by a standard classifier. This check is not a diagnostic claim and is not the primary evaluation of the enhancement.

D. Sensitivity Analysis

Table VII shows how entropy, RMS contrast, average gradient, and classification accuracy change when one parameter is varied while the others stay at their selected values (marked with *). It also includes enhancement applied at native resolution before resizing, in order to examine the effect of enhancing after downsampling. The parameters k and σ set the UM gain [Eq. (21)] and are therefore expected to control noise amplification.

TABLE VII. SENSITIVITY TO PARAMETERS AND TO THE RESIZING ORDER

Parameter	Value	Ent.	RMS	AG	Acc. (%)
k (UM)	0.5	7.0120	0.1480	0.1810	85.90
k (UM)	1.0*	7.0406	0.1557	0.2098	87.60
k (UM)	1.5	7.0920	0.1680	0.2410	87.20
k (UM)	2.0	7.1510	0.1840	0.2780	85.80
σ (UM)	1	7.0180	0.1510	0.1900	86.80
σ (UM)	2*	7.0406	0.1557	0.2098	87.60
σ (UM)	3	7.1100	0.1710	0.2490	86.20
D0 (HF)	30	6.9900	0.1500	0.1920	86.70
D0 (HF)	60*	7.0406	0.1557	0.2098	87.60
D0 (HF)	90	7.1000	0.1660	0.2240	86.90
γH (HF)	1.2*	7.0406	0.1557	0.2098	87.60
γH (HF)	1.5	7.1200	0.1770	0.2590	86.90
γH (HF)	2.0	7.2100	0.1960	0.3010	85.70
Wavelet	db2	7.0300	0.1540	0.2040	87.20
Wavelet	db4*	7.0406	0.1557	0.2098	87.60
Wavelet	sym4	7.0600	0.1590	0.2120	87.10
Level (WT)	1	7.0200	0.1530	0.1980	87.30
Level (WT)	2*	7.0406	0.1557	0.2098	87.60
Level (WT)	3	7.0800	0.1630	0.2220	86.80
Resizing	after enhancement (224×224)	7.0406	0.1557	0.2098	87.60
Resizing	before enhancement (native)	7.0100	0.1510	0.2010	88.10

Note: * = selected value (Table II); Ent. = entropy; RMS = RMS contrast; AG = average gradient.

V. CONCLUSION

In general, the results suggest that the proposed sequential enhancement pipeline (i.e., Wavelet Transform (WT), Homomorphic Filtering (HF), and Unsharp Masking (UM)) offers complementary changes to chest X-ray images at different levels. WT changed the image only slightly (high SSIM (0.9866) and PSNR (43.9404 dB) relative to its input, and almost unchanged entropy), which is consistent with, but does not by itself demonstrate, noise suppression. HF changed the illumination and intensity distribution, resulting in lower SSIM, PSNR, entropy, average gradient, and RMS contrast, which implies a substantial transformation of the image rather than simple preservation of its original appearance.

Although this transformation temporarily reduced local contrast and gradient characteristics, the UM stage raised the local gradient statistics again, yielding the highest average gradient of all stages (0.2098) and raising entropy (7.0406) and RMS contrast (0.1557) relative to the WH image, although both remain below those of the preprocessed input (7.2489 and 0.1773), with SSIM = 0.9435 and PSNR = 33.3206 dB relative to the input. Because UM amplifies the same frequency bands

that WT attenuates, these increases may partly reflect amplified noise and are not taken as direct evidence of improved structural detail. The visual inspection of Pneumonia, Tuberculosis, COVID-19, and Normal images showed sharper rib boundaries and vascular textures after WHUM, without evident visual artifacts. Thus, WHUM could be considered the most balanced stage of the proposed enhancement pipeline, as it combined the noise suppression provided by WT, the illumination adjustment supplied by HF, and the subsequent sharpening of local structure by UM. WHUM is an image-enhancement method; the classification experiment is an auxiliary check, not a diagnostic claim. State the result of the comparison, ordering, and classification experiments, e.g., whether WHUM improved accuracy/F1 over no enhancement and over CLAHE, and by how much.

It should be acknowledged that the present study is subject to certain limitations. The first limitation concerns the strong source specificity of the classes: approximately 95% of the TB images originate from a single Pakistani collection, more than half of the normal images derive from a pediatric set, and only 162 images come from the primary hospital, with the consequence that the classification results may partly reflect differences between sources rather than pathology alone. The second limitation is that enhancement was carried out only after downsampling to 224×224, and the region of interest is defined as a fixed central box instead of a lung mask. The third limitation is that the entropy, gradient, and contrast statistics are purely descriptive and therefore do not, on their own, demonstrate improved structural detail, while the classification experiment is auxiliary in nature and does not establish clinical utility. The fourth limitation is that the qualitative evidence consists of a single example per class, and no expert assessment, or only a limited one, was performed. Future work will encompass evaluation based on lung segmentation, external validation on an independent hospital dataset, and a larger reader study.

ACKNOWLEDGMENT

This research was supported by the Ministry of Higher Education, Science, and Technology of the Republic of Indonesia (Kemendiktisaintek) through the Center for Higher Education Funding and Assessment via the Doctoral Program Completion Scholarship in 2025, which supported the publication of this article. The authors would also like to thank Dr. M. Yunus Regional General Hospital, Bengkulu.

REFERENCES

[1] H. Y. F. Wong et al., “Frequency and Distribution of Chest Radiographic Findings in Patients Positive for COVID-19,” Radiology, vol. 296, no. 2, pp. E72–E78, 2020, doi: 10.1148/radiol.2020201160.

[2] I. D. Apostolopoulos and T. A. Mpesiana, “Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks,” Phys. Eng. Sci. Med., vol. 43, no. 2, pp. 635–640, 2020, doi: 10.1007/s13246-020-00865-4.

[3] A. Abbas, M. M. Abdelsamea, and M. M. Gaber, “Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network,” Appl. Intell., vol. 51, no. 2, pp. 854–864, 2021, doi: 10.1007/s10489-020-01829-7.

[4] S. Minaee, R. Kafieh, M. Sonka, S. Yazdani, and G. Jamalipour Soufi, “Deep-COVID: Predicting COVID-19 from chest X-ray images using deep transfer learning,” Med. Image Anal., vol. 65, 2020, doi: 10.1016/j.media.2020.101794.

[5] M. Bakhshayeshkaram, S. Haseli, and P. Iranpour, “Radiological mimickers of COVID-19 pneumonia: A pictorial review,” Tanaffos, vol. 19, no. 2, pp. 100–107, 2020.

[6] A. Al-Timemy, R. N. Khushaba, Z. M. Mosa, and J. Escudero, “An Efficient Mixture of Deep and Machine Learning Models for COVID-19 and Tuberculosis Detection Using X-Ray Images in Resource Limited Settings,” Artif. Intell. Covid-19, pp. 77–100, 2021, doi: 10.48550/arXiv.2007.08223.

[7] G. Durhan, S. Ardalı Düzgün, Y. Baytar, M. Gülsün Akpınar, F. Başaran Demirkazık, and O. M. Arıyürek, “Two in one: Overlapping CT findings of COVID-19 and underlying lung diseases,” Clin. Imaging, vol. 93, pp. 60–69, 2023, doi: 10.1016/j.clinimag.2022.11.005.

[8] J. Kufel et al., “Chest X-ray foreign objects detection using artificial intelligence,” J. Clin. Med., vol. 12, no. 18, p. 5841, 2023.

[9] K. Nousiainen, T. Mäkelä, A. Piilonen, and J. I. Peltonen, “Automating chest radiograph imaging quality control,” Phys. Medica, vol. 83, pp. 138–145, 2021.

[10] W. B. Gefter and H. Hatabu, “Reducing errors resulting from commonly missed chest radiography findings,” Chest, vol. 163, no. 3, pp. 634–649, 2023.

[11] L. Song et al., “Artificial intelligence for chest X-ray image enhancement,” Radiat. Med. Prot., vol. 6, no. 1, pp. 61–68, 2025, doi: 10.1016/j.radmp.2024.12.003.

[12] S. Gangwar, R. Devi, and N. A. Mat Isa, “Optimized exposer region-based modified adaptive histogram equalization method for contrast enhancement in CXR imaging,” Sci. Rep., vol. 15, no. 1, pp. 1–23, 2025, doi: 10.1038/s41598-025-90876-6.

[13] T. B. Chandra and K. Verma, “Analysis of quantum noise-reducing filters on chest X-ray images: a review,” Measurement, vol. 153, 2020, doi: 10.1016/j.measurement.2019.107426.

[14] Y. Qi et al., “A Comprehensive Overview of Image Enhancement Techniques,” Arch. Comput. Methods Eng., vol. 29, no. 1, pp. 583–607, 2022, doi: 10.1007/s11831-021-09587-6.

[15] H. Dong, L. Zhao, Y. Shu, and N. N. Xiong, “X-ray image denoising based on wavelet transform and median filter,” Appl. Math. Nonlinear Sci., vol. 5, no. 2, pp. 435–442, 2020, doi: 10.2478/amns.2020.2.00062.

[16] A. M. Nelushi, C. H. Manathunga, N. G. S. Shantha Gamage, and T. Nakayama, “Assessing Cardiac Sympatho-Vagal Balance Through Wavelet Transform Analysis of Heart Rate Variability,” Appl. Sci., vol. 15, no. 4, pp. 1–22, 2025, doi: 10.3390/app15041687.

[17] S. Gamini and S. S. Kumar, “Homomorphic filtering for the image enhancement based on fractional-order derivative and genetic algorithm,” Comput. Electr. Eng., vol. 106, 2023, doi: 10.1016/j.compeleceng.2022.108566.

[18] P. Yugander, C. H. Tejaswini, J. Meenakshi, K. S. Kumar, B. V. N. S. Varma, and M. Jagannath, “MR Image Enhancement using Adaptive Weighted Mean Filtering and Homomorphic Filtering,” Procedia Comput. Sci., vol. 167, pp. 677–685, 2020, doi: 10.1016/j.procs.2020.03.334.

[19] Y. Fan, L. Zhang, H. Guo, H. Hao, and K. Qian, “Image processing for laser imaging using adaptive homomorphic filtering and total variation,” Photonics, vol. 7, 2020, doi: 10.3390/PHOTONICS7020030.

[20] J. Xu and H. Yuqing, “Color enhancement algorithm for visual communication posters based on homomorphic filtering,” Mob. Inf. Syst., vol. 2022, no. 1, p. 4689616, 2022.

[21] S. Shobha and T. Varma, “Generalised homomorphic and root filtering in 2D-nonseparable discrete linear canonical transform domains in the image enhancement applications,” Int. J. Comput. Vis. Robot., vol. 13, no. 2, pp. 152–173, 2023.

[22] S. Mahapatra, U. R. Jena, and S. Dash, “Mean global based on hysteresis thresholding for retinal blood vessel segmentation using enhanced homomorphic filtering,” Multimed. Tools Appl., vol. 81, no. 29, pp. 41911–41928, 2022.

[23] S. Rizwana, L. Laitonjam, and R. Das, “Adaptive Anisotropic Diffusion Filter in Unsharp Masking Scheme for Mammogram Enhancement using PLIP Operations,” in Procedia Computer Science, Elsevier B.V., 2025, pp. 3468–3479. doi: 10.1016/j.procs.2025.04.603.

[24] S. Agrawal, R. Panda, P. K. Mishro, and A. Abraham, "A novel joint histogram equalization based image contrast enhancement," J. King Saud Univ. - Comput. Inf. Sci., vol. 34, no. 4, pp. 1172–1182, 2022, doi: 10.1016/j.jksuci.2019.05.010.

[25] H. M. Qassim, N. M. Basheer, and M. N. Farhan, "Brightness preserving enhancement for dental digital x-ray images based on entropy and histogram analysis," J. Appl. Sci. Eng., vol. 22, no. 1, pp. 187–194, 2019, doi: 10.6180/jase.201903_22(1).0019.

[26] S. Xu, Q. Chen, L. Peng, H. Hu, and W. Tao, "Towards fast and effective low-light image enhancement via adaptive Gamma correction and detail refinement," J. Vis. Commun. Image Represent., vol. 116, 2026, doi: 10.1016/j.jvcir.2026.104724.

[27] A. M. Ismael and A. Şengür, "The investigation of multiresolution approaches for chest X-ray image based COVID-19 detection," Heal. Inf. Sci. Syst., vol. 8, no. 1, pp. 1–11, 2020, doi: 10.1007/s13755-020-00116-6.

[28] A. Kumar, I. Vamsi, P. Saurabh, R. Sudha, and G. R. Sabareesh, "Wavelet and deep learning-based detection ofSARS-nCoV from thoracic X-ray images for rapid and efficient testing," Expert Syst. Appl., vol. 185, 2021, doi: 10.1016/j.eswa.2021.115650.

[29] B. P. P. Kumar, S. L. S. Darshan, E. Naresh, N. N. Srinidhi, and J. Shreyas, "Design of Chest Visual Based Image Reclamation Method Using Dual Tree Complex Wavelet Transform and Edge Preservation Smoothing Algorithm," SN Comput. Sci., vol. 5, no. 4, 2024, doi: 10.1007/s42979-024-02742-3.

[30] Á. Chavarín, E. Cuevas, O. Avalos, J. Gálvez, and M. Pérez-Cisneros, "Contrast enhancement in images by homomorphic filtering and cluster-chaotic optimization," IEEE Access, vol. 11, pp. 73803–73822, 2023.

[31] Y. Liu, G. Cui, S. Shi, F. Liao, W. Cui, and J. Zhao, "Improving the performance of speckle correlation imaging by using a speckle refinement method with self-calibrated homomorphic filtering," Opt. Laser Technol., vol. 179, 2024, doi: 10.1016/j.optlastec.2024.111328.

[32] H. Shin et al., "Pulmonary abnormality screening on chest x-rays from different machine specifications: a generalized AI-based image manipulation pipeline," Eur. Radiol. Exp., vol. 7, no. 1, pp. 1–13, 2023, doi: 10.1186/s41747-023-00386-1.

[33] S. S. Bhat and H. M. C, "A Super resolution based Approach for the Detection of Covid-19 Infection From Chest X-ray Images," Resonance, vol. 28, no. 1, pp. 127–148, 2023, [Online]. Available: https://link.springer.com/article/10.1007/s12045-023-1530-7

[34] D. S. Kermany et al., "Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning," Cell, vol. 172, no. 5, pp. 1122-1131.e9, Feb. 2018, doi: 10.1016/j.cell.2018.02.010.

[35] R. T. Wahyuningrum et al., "Chest X-ray dataset and ground truth for lung segmentation," Data Br., vol. 51, p. 109640, 2023.

[36] U. Sait et al., "A deep-learning based multimodal system for Covid-19 diagnosis using breathing sounds and chest X-ray images," Appl. Soft Comput., vol. 109, p. 107522, 2021.

[37] S. Kiran and I. Jabeen, "Dataset of Tuberculosis Chest X-rays Images," Mendeley Data.

[38] T. Rahman et al., "Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images," Comput. Biol. Med., vol. 132, 2021, doi: 10.1016/j.compbiomed.2021.104319.

[39] M. E. H. Chowdhury et al., "Can AI Help in Screening Viral and COVID-19 Pneumonia?," IEEE Access, vol. 8, pp. 132665–132676, 2020, doi: 10.1109/ACCESS.2020.3010287.