

Interpretable Radiomic Analysis of Chest X-rays for Pneumonia Discrimination Using First-Order and GLCM Texture Features, Statistical Screening, and PCA

Muhammad Noman

2024msbme6@student.uet.edu.pk

University of Engineering and Technology Lahore <https://orcid.org/0009-0009-5576-2671>

Muhammad Aziz

University of Lahore

Farooq Mukhtar

University of Engineering and Technology Lahore

Research Article

Keywords: radiomics, chest X-ray, pneumonia, GLCM, texture analysis, U-Net, principal component analysis, interpretable machine learning

Posted Date: October 22nd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11447618/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Abstract

Background

Chest X-ray (CXR) is the first-line imaging test for respiratory disease, yet its interpretation is subjective and varies between and within readers. Radiomics offers an objective, quantitative alternative by extracting measurable descriptors from image regions.

Methods

We built a transparent radiomics pipeline on 2,000 images from the NIH ChestX-ray14 dataset (1,000 "No Finding", 1,000 labelled pneumonia). Images were contrast-enhanced with CLAHE and the lungs were segmented with a pre-trained U-Net; 89 images with corrupted masks were excluded (final N = 1,911). Eleven features were extracted from the segmented lung region: five first-order statistics (mean, standard deviation, skewness, kurtosis, entropy) and six gray-level co-occurrence matrix (GLCM) texture features (contrast, dissimilarity, homogeneity, energy, correlation, angular second moment). Features were screened with independent-samples t-tests, standardized, and reduced with principal component analysis (PCA). Logistic regression, support vector machine (SVM) and random forest classifiers were evaluated with stratified 5-fold cross-validation.

Results

Ten of the eleven features differed significantly between classes ($p < 0.05$); only mean intensity did not ($p = 0.313$). Effect sizes were small to moderate (approximate Cohen's d up to 0.54), with the largest differences in GLCM energy, entropy, ASM, homogeneity, dissimilarity and contrast. Four principal components retained 90.94% of the variance, with the first (62.02%) dominated by GLCM texture-uniformity features. Classification performance was moderate: the SVM was best (accuracy 0.637, AUC 0.687), with logistic regression and random forest close behind.

Conclusions

Texture-based radiomic features carry statistically significant, interpretable information that separates pneumonia from normal CXRs, whereas global brightness does not. Moderate classification accuracy indicates that such a compact feature set is best positioned as a supporting, explainable tool rather than a stand-alone diagnostic, and that richer higher-order and localized features are needed.

1. Introduction

Respiratory diseases such as pneumonia and lung cancer remain a serious burden on health systems worldwide, and the COVID-19 pandemic underscored the importance of rapid identification of lung

pathology. Chest X-ray (CXR) is among the most frequently performed imaging examinations. It is inexpensive, quick, available at the bedside through portable units, and is often the only imaging modality available in resource-constrained settings. Compared with computed tomography (CT), it involves a lower radiation dose and avoids long decontamination intervals between patients [1], [2].

Despite this role, CXR interpretation relies largely on qualitative visual inspection. Diagnoses may differ between observers and for the same observer at different times, and disagreement rates of up to 50% have been reported in some situations [3]. Many lung diseases share similar radiographic appearances; for example, the bilateral infiltrates and ground-glass opacities of COVID-19 resemble those of SARS and MERS [2], [3]. In addition, a CXR is a two-dimensional projection of a three-dimensional structure, so ribs, clavicles and the cardiac silhouette can overlap with, or mimic, lung abnormalities [1], [4]. Early changes are subtle and are often recognized only by experienced radiologists.

Radiomics addresses this limitation by treating images as data: quantitative features describing the intensity distribution and spatial texture of a region of interest (ROI) are extracted by computation and linked to the underlying tissue pathophysiology [4], [5]. Features are commonly grouped into first-order statistics (e.g. skewness, kurtosis), second-order texture features such as those derived from the gray-level co-occurrence matrix (GLCM), and higher-order features obtained after image filtering [5]. Computer-aided diagnosis systems use such features to train classifiers including support vector machines (SVM), random forests (RF) and k-nearest neighbours. Compared with deep convolutional networks, which behave as "black boxes", hand-crafted radiomic features offer a more transparent framework in which each input corresponds to a measurable image property [6], [7].

In this study we develop and evaluate a compact, fully interpretable radiomics pipeline for distinguishing pneumonia from normal CXRs. Our contributions are:

1. A reproducible pipeline combining CLAHE enhancement, deep-learning lung segmentation and extraction of a compact set of eleven first-order and GLCM features from the segmented lungs.
2. A statistical characterization, on a balanced cohort of 1,911 images, of which features discriminate the two classes and by how much, including approximate effect sizes.
3. A PCA-based analysis that explains the redundancy and latent structure of the feature set in physically meaningful terms.
4. A baseline classification benchmark (LR, SVM, RF) under stratified cross-validation, with an explicit discussion of limitations and the steps needed to improve performance.

2. Related Work

Histogram and texture analysis of CXR. Gray-level histogram patterns differ between normal and diseased lungs: normal lungs tend to show a broader intensity distribution with more brightness variation, whereas pneumonia and COVID-19 images show a narrower, skewed distribution consistent with a more "cloudy" appearance [8]. GLCM-based texture descriptors have been used with classical and

deep classifiers to separate lung disease classes [9], and texture features have been applied to COVID-19 pneumonia on CXR in proof-of-concept and classification studies [2], [10].

Feature selection and dimensionality reduction. Radiomic data are high-dimensional, so feature screening is usually applied before modelling to reduce the risk of fitting noise. Student's t-tests, Wilcoxon rank-sum tests and one-way ANOVA are common filters. Manav et al. combined feature selection with machine learning on CXR [11], and Tamal et al. proposed an integrated radiomics and machine-learning framework for early COVID-19 diagnosis on CXR [12]. Khan et al. ranked GLCM features and identified homogeneity as especially informative for infection classes [13]. PCA is widely used to summarize correlated texture features compactly, and dimensionality reduction has been reported to improve classification performance [4].

Segmentation. Classical thresholding, graph-cut and active-contour methods struggle with the complex gray-level patterns of pulmonary tissue, whereas U-Net-based networks learn hierarchical features and produce accurate lung masks even in the presence of devices and leads [5], [14]. Restricting feature extraction to anatomically relevant lung regions is therefore an important step in any CXR radiomics workflow.

Gap. Many published radiomic CXR pipelines use very large feature sets or opaque models. There is value in a minimal, well-documented baseline whose every feature can be interpreted clinically, and in quantifying how far such a baseline can go. This work provides that baseline.

3. Materials and Methods

3.1 Dataset

Images were drawn from the NIH ChestX-ray14 dataset, which contains 112,120 frontal chest radiographs from 30,805 unique patients in 8-bit grayscale format, with anonymized metadata [15]. Labels in this dataset were generated from radiology reports by natural language processing (NLP) [15].

A balanced subset of $N = 2,000$ posteroanterior (PA) radiographs was sampled with a custom Python script (random seed 42): 1,000 images labelled "No Finding" (normal/control) and 1,000 images whose finding labels include "Pneumonia" (pathology group). Cases in which pneumonia co-occurred with other conditions were retained rather than excluded, so the pathology class reflects realistic multi-morbidity.

3.2 Preprocessing

All images were loaded, converted to 8-bit grayscale and resized to 256×256 pixels for network input. To reduce bias from differing acquisition parameters, contrast was standardized with contrast-limited adaptive histogram equalization (CLAHE), which divides the image into small non-overlapping tiles (8×8 grid) and equalizes each locally, with a clip limit of 2.0 to constrain contrast amplification.

3.3 Lung Segmentation

Lungs were segmented with a standard four-level U-Net with skip connections implemented in TensorFlow/Keras (encoder filters 32–64–128–256, a 512-filter bottleneck, and a symmetric decoder using transposed convolutions, ending in a 1×1 sigmoid output). The network was initialized with publicly available pre-trained weights from a lung-segmentation notebook [16]; no additional training was performed. The output probability map was thresholded at 0.5 to obtain a binary mask, which was resized to the original image dimensions with nearest-neighbour interpolation and applied to the original grayscale image by a bitwise AND. Images whose masks were corrupted, typically because lung boundaries were obscured, were excluded after inspection, leaving $N = 1,911$ images (955 normal, 956 pneumonia).

3.4 Radiomic Feature Extraction

Features were computed only on non-zero pixels inside the segmented lung region. Eleven features were extracted.

First-order statistics (5). Mean, standard deviation, skewness, kurtosis and entropy of the ROI intensity values. Entropy was computed as the sum over ROI pixels of $-(I/255) \cdot \log_2(I/255 + \epsilon)$, with I the pixel intensity and $\epsilon = 10^{-10}$; note that, as implemented, this quantity grows with the number of ROI pixels.

Second-order GLCM texture features (6). A symmetric, normalized GLCM was computed at 256 gray levels with pixel distance $d = 1$ and angles of 0° , 45° , 90° and 135° . Contrast, dissimilarity, homogeneity, energy, correlation and angular second moment (ASM) were derived from the matrix, and each was averaged over the four directions.

3.5 Statistical Analysis and Feature Optimization

Feature screening. Each feature was compared between classes with an independent-samples t-test (two-sided, $\alpha = 0.05$). Levene's test selected between the pooled-variance Student's t-test and Welch's t-test for each feature. Features with $p < 0.05$ were retained. Because eleven tests were performed, we also checked the results against a Bonferroni-adjusted threshold ($0.05/11 \approx 0.0045$). Approximate standardized effect sizes (Cohen's d) were derived from the t-statistics and group sizes as $d \approx t \cdot \sqrt{(1/n_1 + 1/n_2)}$.

Normalization. Retained features were standardized by z-score scaling.

Dimensionality reduction. PCA was applied to the standardized features. The number of components was chosen as the smallest number whose cumulative explained-variance ratio reached at least 90%. Component loadings were examined to interpret each component.

3.6 Classification and Evaluation

Three baseline classifiers were evaluated: logistic regression (LR) as a linear baseline, a support vector machine (SVM) with probability estimation operating in a higher-dimensional kernel space, and a random forest (RF) to capture non-linear relationships. The classifiers were trained on the four principal

components retained from the PCA described in Section 3.5. All models were implemented in scikit-learn with default settings (LR: L2 penalty, $C = 1.0$; SVM: radial basis function kernel, $C = 1.0$, $\gamma = \text{"scale"}$; RF: 100 trees, Gini criterion). The pipeline was run in Python (TensorFlow/Keras, OpenCV, scikit-image, SciPy and scikit-learn). Performance was estimated with stratified 5-fold cross-validation, so that the class balance was preserved in every fold. Z-score standardization and PCA were fitted without using class labels, on the full feature matrix, before cross-validation; this unsupervised step is acknowledged in the limitations. We report accuracy, precision, recall, F1-score, specificity and area under the ROC curve (AUC), with pneumonia as the positive class. Because the three models were run with default settings and no hyperparameter tuning, the comparison reflects baseline performance rather than the best attainable for each model.

4. Results

4.1 Cohort and Segmentation

Of the 2,000 sampled images, 89 were removed because of corrupted masks, giving a final cohort of 1,911 images (955 normal, 956 pneumonia). Qualitatively, the pre-trained U-Net produced clean binary masks that isolate the lung fields and exclude most bone and extrathoracic structures (Fig. 1).

4.2 Feature Significance

Ten of the eleven features differed significantly between normal and pneumonia images (Table 1). Mean intensity was the only non-significant feature ($p = 0.313$, $d \approx 0.05$) and was therefore excluded from further analysis. All ten remaining features also survive the Bonferroni-adjusted threshold (largest $p = 0.0016$ for correlation). Because the cohort is large, significance should be read together with effect size: the first-order standard deviation, skewness, kurtosis and the GLCM correlation showed small effects ($|d| < 0.3$), while entropy, contrast, dissimilarity, homogeneity, energy and ASM showed moderate effects ($|d| \approx 0.49\text{--}0.54$).

Table 1

Independent-samples t-test results between normal (n = 955) and pneumonia (n = 956) groups. Cohen's d is approximate, derived from t; positive values indicate higher values in the normal class.

Feature	Mean (Normal)	Mean (Pneumonia)	t	p	Approx. d	Significant
Mean	92.984	91.832	1.009	0.31331	0.05	No
Standard deviation	26.884	25.643	3.862	0.00012	0.18	Yes
Skewness	0.559	0.434	5.861	< 0.00001	0.27	Yes
Kurtosis	0.220	0.030	4.157	0.00003	0.19	Yes
Entropy	94609.12	72901.69	11.844	< 0.00001	0.54	Yes
Contrast	41.656	32.198	10.641	< 0.00001	0.49	Yes
Dissimilarity	0.767	0.603	11.018	< 0.00001	0.50	Yes
Homogeneity	0.875	0.903	-11.259	< 0.00001	-0.52	Yes
Energy	0.813	0.856	-11.773	< 0.00001	-0.54	Yes
Correlation	0.984	0.982	3.156	0.00163	0.14	Yes
ASM	0.667	0.739	-11.718	< 0.00001	-0.54	Yes

The direction of the effects is coherent. Normal lungs showed higher entropy, contrast, dissimilarity, standard deviation, skewness and kurtosis, whereas pneumonia lungs showed higher homogeneity, energy and ASM (Fig. 2).

The correlation heatmap revealed strong redundancy among the GLCM features. Homogeneity, energy and ASM were almost perfectly correlated with each other ($r \approx 0.99-1.00$) and strongly anti-correlated with entropy and dissimilarity (r from -0.96 to -1.00), indicating that they carry largely the same information (Fig. 3).

4.3 Principal Component Analysis

PCA of the ten standardized features required four components to exceed the 90% variance target; together they explained 90.94% of the total variance, and the first component alone explained 62.02%. The loadings (Table 2) give each component a clear physical meaning:

- **PC1 (texture uniformity)**: high positive loadings for homogeneity, energy and ASM ($\approx +0.39$ to $+0.40$) and high negative loadings for entropy, dissimilarity and contrast (≈ -0.33 to -0.39). PC1 therefore contrasts uniform, repetitive texture against heterogeneous texture.
- **PC2 (histogram shape)**: driven by kurtosis ($+0.635$) and skewness ($+0.505$), opposed by correlation (-0.541).
- **PC3 (intensity spread)**: almost exclusively standard deviation ($+0.911$), independent of spatial texture.
- **PC4 (linear pixel dependency)**: dominated by correlation ($+0.816$) with a contribution from skewness ($+0.394$).

Table 2
PCA component loadings of the ten retained features.

Feature	PC1	PC2	PC3	PC4
Standard deviation	-0.241	-0.040	0.911	-0.110
Skewness	-0.185	0.505	0.220	0.394
Kurtosis	-0.081	0.635	-0.068	0.290
Entropy	-0.391	-0.018	-0.200	-0.017
Contrast	-0.330	-0.213	0.064	-0.259
Dissimilarity	-0.394	0.002	-0.060	-0.109
Homogeneity	0.395	-0.038	0.140	0.050
Energy	0.394	0.002	0.165	0.034
Correlation	-0.126	-0.541	0.032	0.816
ASM	0.395	0.010	0.147	0.011

In the two-dimensional PC1–PC2 projection, normal and pneumonia samples overlapped substantially and did not form distinct clusters (Fig. 4), which anticipates the moderate classification performance reported below.

4.4 Classification Performance

All three classifiers achieved moderate, similar performance under stratified 5-fold cross-validation (Table 3). The SVM had the highest accuracy (0.637), specificity (0.718), precision (0.652) and AUC (0.687); random forest had the highest recall (0.582); logistic regression was between the two on most metrics.

Table 3
Cross-validated classification performance (pneumonia = positive class).

Model	Accuracy	Precision	Recall	F1-score	Specificity	AUC
Logistic regression	0.630	0.634	0.573	0.601	0.684	0.672
SVM	0.637	0.652	0.553	0.597	0.718	0.687
Random forest	0.622	0.620	0.582	0.600	0.661	0.661

5. Discussion

Texture, not brightness, carries the signal. The non-significance of mean intensity ($p = 0.313$) shows that overall lung brightness is too similar between classes to be informative, consistent with prior reports that mean intensity of COVID-19 and normal CXRs can overlap [8]. In contrast, all six GLCM features differed significantly with moderate effects, indicating that the spatial arrangement of pixel intensities is the more informative characteristic.

Physical interpretation. Aerated lung produces a heterogeneous radiographic pattern from vessels and airway walls against air-filled parenchyma, giving higher contrast, dissimilarity and entropy. Pneumonia replaces air in alveolar spaces with fluid and cellular material, producing consolidation and ground-glass opacity that raise attenuation and flatten local variation. This is consistent with the observed higher homogeneity, energy and ASM, and with the lower standard deviation, contrast and dissimilarity in the pneumonia group, and with the narrower intensity distributions described for diseased lungs in earlier work [8]. Principal component analysis expresses this finding compactly: PC1, a texture-uniformity axis, captures 62% of the variance and is built from exactly the features with the largest effect sizes. PC3 isolates the global intensity spread, which is independent of spatial texture.

Redundancy. Homogeneity, energy and ASM, and likewise entropy and dissimilarity, are nearly collinear. The strong significance of ten features should therefore not be read as ten independent findings; PCA shows that the information content is closer to a few latent factors, and a practical pipeline could drop redundant GLCM features without losing information.

Moderate classification performance. An AUC of about 0.69 and accuracy of about 0.64 show that the features contain meaningful but incomplete discriminative information. Large-sample significance (t-tests with about 1,900 images) does not imply strong per-image separability: moderate effect sizes ($d \approx 0.5$) and heavily overlapping distributions (Figs. 2 and 4) limit what any classifier can achieve. Reasons include the small, global, hand-crafted feature set; the inclusion of pneumonia cases co-occurring with other pathologies; label noise; and the absence of higher-order (GLRLM, GLSZM, wavelet) features, which earlier studies have reported to improve performance [5], [12]. Differences among LR, SVM and RF were small, which suggests the limiting factor is the information in the features rather than classifier complexity.

Interpretability. Despite the moderate accuracy, every input to the model is a measurable image property with a clear meaning, so a clinician can see why a case is flagged. This makes the approach suitable as a supporting decision tool alongside, not instead of, expert reading, in line with the case made for feature-based models over black-box deep networks [6], [7].

6. Limitations

1. **Label noise.** ChestX-ray14 labels were extracted by NLP from reports, and the ground truth may contain an error rate of up to about 10%. Pneumonia cases were retained with co-occurring conditions, and "normal" means "No Finding".
2. **Global features only.** Without lesion-level annotations, features were computed over the entire segmented lung area, which dilutes focal abnormalities.
3. **Segmentation dependence and validation.** The pipeline relies on a pre-trained U-Net that was not retrained or quantitatively validated on this cohort (no Dice or IoU against reference masks). Masks fail when lung boundaries are obscured, as in severe effusion, massive consolidation or white-out lungs; the 89 excluded images were removed on this basis, which may bias the cohort toward easier cases.
4. **Feature set.** Only eleven first-order and GLCM features were used, several of them highly redundant, and entropy as implemented scales with ROI size.
5. **Evaluation.** Performance was estimated by internal cross-validation on a single dataset; no external validation was performed. Because ChestX-ray14 contains multiple images per patient, cross-validation should be grouped by patient to rule out optimistic bias. In addition, standardization and PCA were fitted on the full cohort (without class labels) before cross-validation rather than inside each training fold, which may introduce a small optimistic bias; fold-wise fitting is recommended in future work.
6. **Task scope.** The study addresses a binary normal-versus-pneumonia task rather than the full spectrum of lung disease.

7. Conclusion and Future Work

This study presents an interpretable radiomics framework for chest X-ray analysis combining CLAHE, deep-learning lung segmentation, first-order and GLCM features, statistical screening and PCA. Ten of eleven features differed significantly between normal and pneumonia images, with texture-uniformity features showing the largest effects, while global mean intensity did not. PCA reduced the features to four components that preserved 90.94% of the variance and carried clear physical meaning. Baseline classifiers reached moderate performance (best AUC 0.687), showing that this compact feature set is informative but insufficient on its own.

Future work should (i) add higher-order features (GLRLM, GLSZM, wavelet and LoG-filtered features); (ii) move from global to localized feature extraction using object detection or lesion annotations; (iii) validate segmentation quantitatively and evaluate patient-level splits and external datasets; (iv) compare

against and combine with deep-learning classifiers while preserving interpretability, for example through hybrid models; and (v) extend the task to multi-class disease discrimination.

Declarations

Ethics

This study used only publicly available, fully anonymized data from the NIH ChestX-ray14 dataset. No human participants were recruited and no identifiable patient information was accessed; therefore, separate ethical approval and informed consent were not required.

Conflicts of interest:

The authors declare that they have no conflict of interest.

Funding:

The authors received no specific funding for this work.

Author contributions:

M.N. (main author): conceptualization, methodology, software, data curation, formal analysis, investigation, visualization, and writing of the original draft. M.A. (co-supervisor): supervision, validation, and writing (review and editing). F.M. (supervisor): supervision, project administration, and writing (review and editing). All authors read and approved the final manuscript.

Acknowledgments:

The authors thank the administrative staff of the Biomedical and Electrical Engineering Departments, University of Engineering and Technology, Lahore.

Data availability:

The source images are publicly available from the NIH ChestX-ray14 dataset. The extracted feature dataset and the analysis code are available from the corresponding author on reasonable request.

References

1. Susanto F, Utami HS, Pradika FR, Idris FF, Febriana C, Kurniawan MY (2023) Analysis of histogram and grayscale on chest X-ray in lung cancer using Image-J. *Indones J Electron Electromed Eng Med Inf* 5(3):181–186
2. Cavallo AU, Troisi J, Forcina M, Mari PV, Forte V, Sperandio M et al (2021) Texture analysis in the evaluation of Covid-19 pneumonia in chest X-ray images: a proof of concept study. *Curr Med Imaging Rev* 17(9):1094–1102
3. Hussain L, Nguyen T, Li H, Abbasi AA, Lone KJ, Zhao Z et al (2020) Machine-learning classification of texture features of portable chest X-ray accurately classifies COVID-19 lung infection. *BioMed Eng OnLine* 19(1):88
4. Ortiz-Toro C, García-Pedrero A, Lillo-Saavedra M, Gonzalo-Martin C (2022) Automatic detection of pneumonia in chest X-ray images using textural features. *Comput Biol Med* 145:105466
5. Ferreira Junior JR, Cardona Cardenas DA, Moreno RA, de Sa Rebelo MDF, Krieger JE, Gutierrez MA (2021) Novel chest radiographic biomarkers for COVID-19 using radiomic features associated with diagnostics and outcomes. *J Digit Imaging* 34(2):297–307
6. Kim YJ (2022) Machine learning model based on radiomic features for differentiation between COVID-19 and pneumonia on chest X-ray. *Sensors* 22(17):6709
7. Zhang Y, Yang D, Lam S, Li B, Teng X, Zhang J et al (2022) Radiomics-based detection of COVID-19 from chest X-ray using interpretable soft label-driven TSK fuzzy classifier. *Diagnostics* 12(11):2613
8. Sumarti H, Sabrina Q, Triana D, Septiani F, Rahmani TPD (2023) Identification of Covid-19 based on features texture histogram and gray level co-occurrence matrix (GLCM) using k-means clustering methods in chest X-ray digital images. *J Penelit Fis Apl* 13(1):51–66
9. Nurcahyati I, Saragih TH, Farmadi A, Kartini D, Muliadi M (2024) Classification of lung disease in X-ray images using gray level co-occurrence matrix method and convolutional neural network. *J Electron Electromed Eng Med Inf* 6(4):332–342
10. Moura L, Dartora C, Mattjie C, Barros R (2021) Texture-based feature extraction for Covid-19 pneumonia classification using chest radiography. *EAI Endorsed Trans Bioeng Bioinform*
11. Manav M, Goyal M, Kumar A (2023) Role of optimal features selection with machine learning algorithms for chest X-ray image analysis. *J Med Phys* 48(2):195–203
12. Tamal M, Alshammari M, Alabdullah M, Hourani R, Alola HA, Hegazi TM (2021) An integrated framework with machine learning and radiomics for accurate and rapid early diagnosis of COVID-19 from chest X-ray. *Expert Syst Appl* 180:115152
13. Khan RNM, Majid A, Shim SO, Habibullah S, Almazroi AA, Hussain L (2024) Intelligent Bayesian inference for multiclass lung infection diagnosis: network analysis of ranked gray level co-occurrence (GLCM) features. *New Gener Comput* 42(5):997–1048
14. Yang D, Ren G, Ni R, Huang YH, Lam NFD, Sun H et al (2023) Deep learning attention-guided radiomics for COVID-19 chest radiograph classification. *Quant Imaging Med Surg* 13(2):572
15. Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM (2017) ChestX-ray8: hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax

diseases. In: Proc IEEE Conf Comput Vis Pattern Recognit (CVPR). :2097–2106

16. Pandey N (2020) Lung segmentation from chest X-ray dataset. Kaggle Noteb.

<https://www.kaggle.com/code/nikhilpandey360/lung-segmentation-from-chest-x-ray-dataset>

Figures

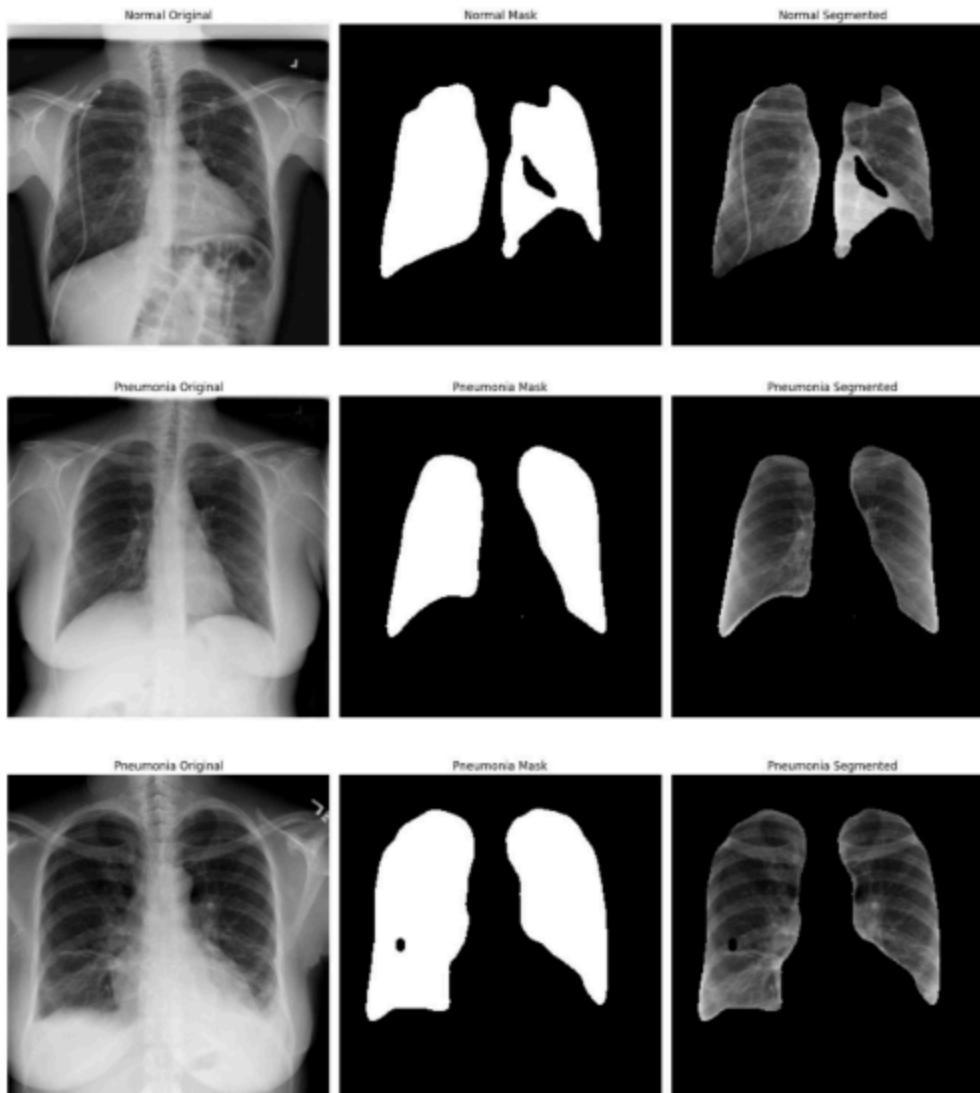


Figure 1

U-Net lung segmentation. Top row: normal CXR; middle and bottom rows: pneumonia CXRs. Columns show the original image (left), the binary lung mask (centre) and the segmented lung region (right).

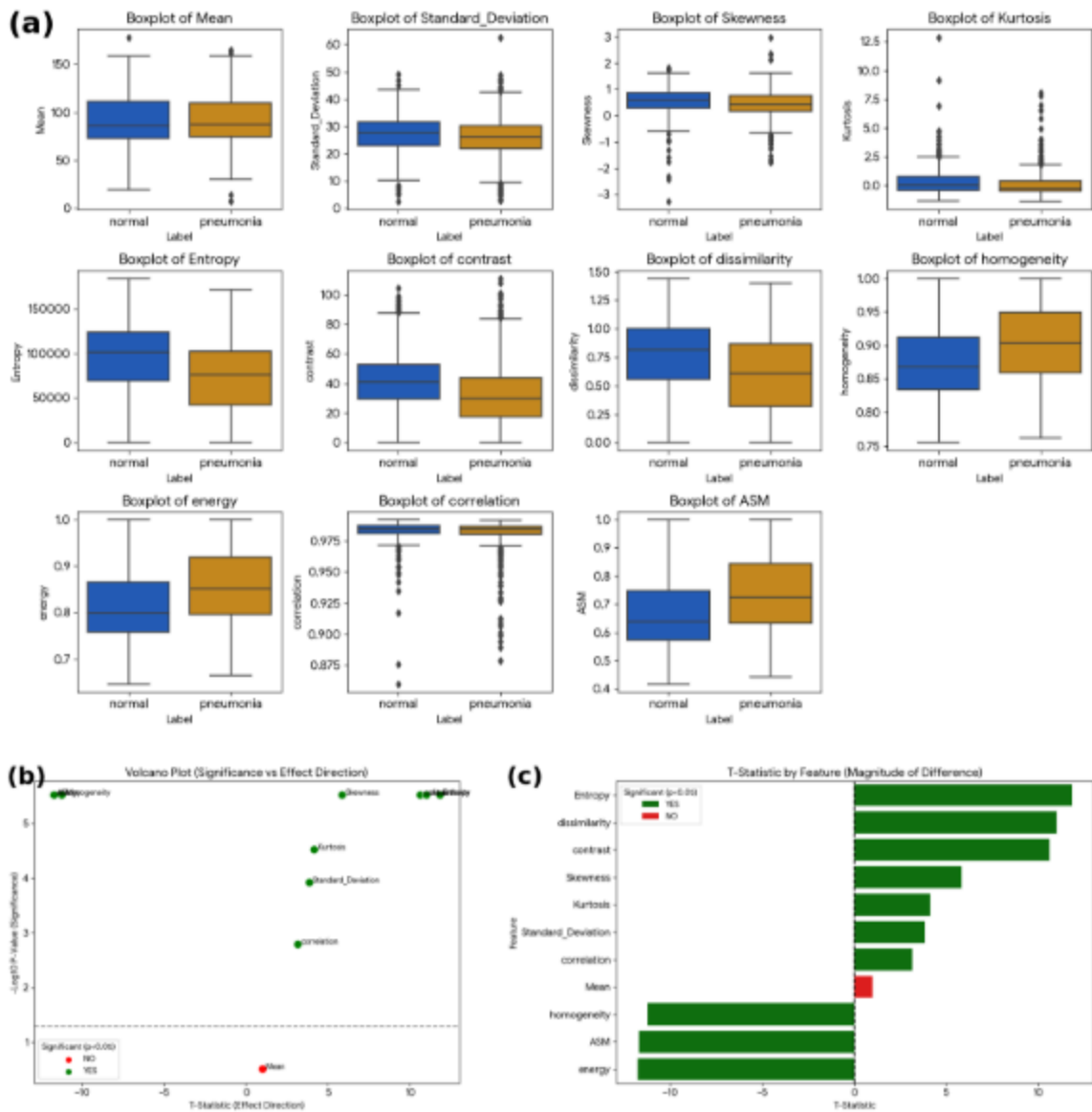


Figure 2

Feature distributions and significance. (a) Boxplots of the eleven features by class; (b) volcano plot (significance versus effect); (c) bar chart of the t-statistic for each feature, showing magnitude and direction.

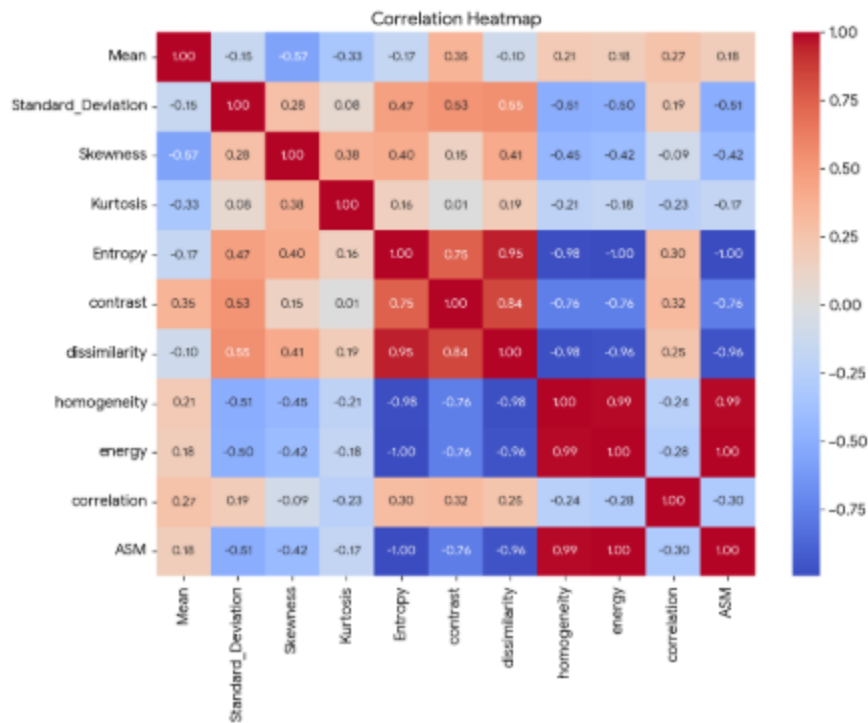


Figure 3

Pearson correlation heatmap of the eleven radiomic features, showing the strong redundancy among the GLCM features.

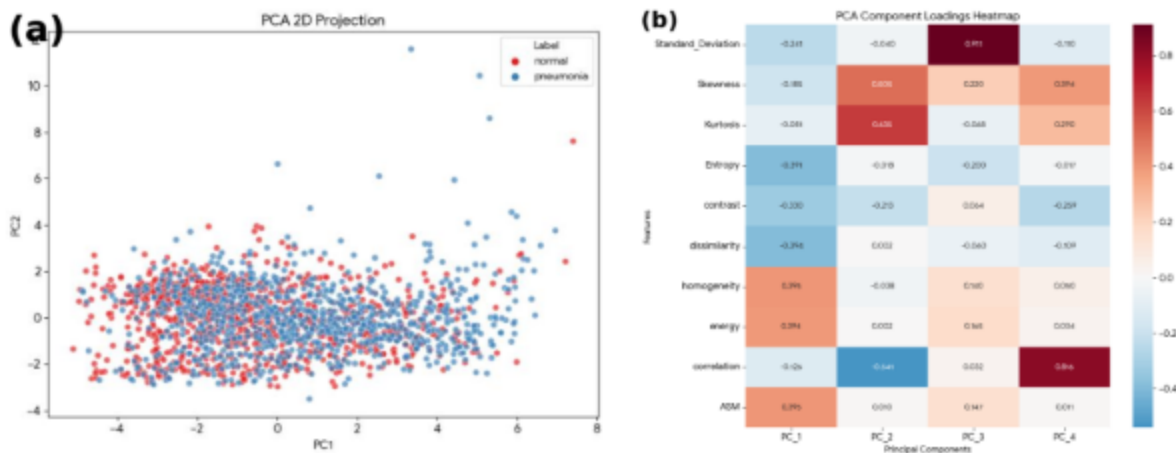


Figure 4

PCA of the ten retained features. (a) PC1–PC2 projection coloured by class, showing substantial overlap between normal and pneumonia images; (b) component-loading heatmap for PC1–PC4.