Methodology

Topic: Preparing high-quality X-ray images for AI-based diagnostic systems

Group: SE - 2307

Names: Saulet Kabdrakhmanov, Galymzhan Aliakbar, Aruzhan Atelova and Bekzhan Nurallin.

Course: Research Methods and Tools

Instructor: Seitenov Altynbek

Table of contents

Research questions

RQ1: How does chest X-ray image quality affect AI model performance?
RQ2: Which preprocessing techniques most effectively improve images before AI analysis?
RQ3: Can a standardized preparation pipeline make AI results more reliable across sites?

Research aim

The aim of this study is to improve the quality and consistency of chest X-ray images used in AI-based diagnostic systems through the evaluation of preprocessing techniques and the development of a standardized image preparation pipeline.

Previous studies have shown that the performance of traditional machine learning models "relies heavily on the quality of the extracted image features," and that these features "are often sensitive to various conditions such as light and the object's orientation within the image, as well as noise and other factors" (Elyan et al., 2022, p. 26). This means that having clear and consistent image data is very important, especially in medical fields where accuracy is critical.

According to Elyan et al. (2022), "data availability and quality play a crucial role in the learning process" (p. 34). In medical imaging, differences in equipment, lighting, and how images are taken can greatly affect how well a model can generalize and make accurate predictions. The authors also mention that "strictly controlled and consistent conditions specifically for data gathering purposes are not always possible," which leads to "challenges in the generalisability of models" (p. 34). Because of this, our research focuses on preprocessing and standardization to help reduce these problems. As they further explain, "a complete lack of standardisation in data gathering protocols in some fields will produce diverse but disjoint datasets, making model generalisation exceptionally difficult" (p. 34). Therefore, creating a standardized preprocessing pipeline for chest X-rays should improve both the quality of the images and the robustness of AI models.

Similarly, Olveres et al. (2023) point out that "the poor quality of ultrasound images due to speckle noise makes the differentiation difficult" (p. 3835). The same issue can happen in chest X-ray images, where changes in contrast, lighting, or artifacts can make it harder for AI models to detect abnormalities. For this reason, a preprocessing step is essential to "obtain more homogenous regions and enhance the contrast of the image, while preserving important diagnostic features" (Olveres et al., 2023, p. 3835). They also note that "the quality of the images will also depend on the device, contributing to the difficulty of modeling the lesions" (p. 3836). This again supports the idea that having a standardized preprocessing process for medical images can help make diagnostic AI systems more consistent and reliable, regardless of differences in devices or imaging conditions.

Liu, Siegel, and Shen (2022) also highlight that "most publicly available COVID-19 imaging data do not retain the source image information on the diagnosed objects… the quality of those shared COVID-19 images is degraded" (pp. 189–190). This shows why proper preprocessing and standardization are urgently needed. They even suggest that "an international consortium of domain experts should be formed to address these COVID-19 imaging data trustworthiness issues in preprocessing, curation, standardization, annotation with consensus as ground truth, and sharing" (p. 189). This directly aligns with the goal of this study—to build a standardized image preparation pipeline that ensures data reliability before AI models are applied.

Moreover, "fair and unbiased deep learning models heavily depend on the availability of high-quality annotated and curated benchmarked data sets" (Liu et al., 2022, p. 188). Projects like the Medical Imaging and Data Resource Center (MIDRC), which "plans to serve as a linked-data commons that coordinates access to data and harmonizes data management activities at three phases: curation, annotation, and quality assessment" (pp. 192–193), show how important it is to maintain consistency and quality in medical imaging data.

This research aims to reduce image degradation and improve the consistency of medical imaging data by applying normalization, denoising, and contrast enhancement techniques. As Elyan et al. (2022) point out, "the quality of the data collected (images, videos) may be unintentionally degraded in an uncontrolled environment, and this, in turn, will have a negative impact on the performance of any DL model" (p. 38). By improving data consistency and quality, this study hopes to make AI systems for chest X-ray analysis more reliable and generalizable. It also supports the broader goal described by Liu et al. (2022), that "the wider availability of high-quality, curated, and benchmarked imaging data sets offers great promise that domain experts in medical imaging and deep learning can collaboratively advance… responsible and trustworthy deep learning" (p. 194).

Research schedule

Week	Research Phase	Objectives / Tasks	Deadline
Week 6	**Data Review and Preparation**	- Collect all required chest X-ray datasets. - Check image quality and identify noise, artefacts, and incorrect labels. - Create a catalog separating "clean" and "noisy" images.	End of Week 6
Week 7	**Data Cleaning and Enhancement**	- Apply preprocessing and	End of Week 7

	& **Writing introduction chapter**	enhancement techniques (e.g., noise reduction, contrast adjustment). - Remove low-quality or duplicate images. - Develop a script for automated data cleaning. -Read and analyze relevant extra articles(if necessery) -Equaly separate task among the team -Discass with team given task and try to find common solution	
Week 8	**Model Testing with Improved Dataset**	- Train a baseline CNN model using the improved dataset. - Compare model accuracy with results from the original dataset. - Analyze how data quality affects model performance.	End of Week 8
Week 9	**Evaluation and Validation** **&** **Writing methodology**	- Evaluate model performance using metrics . - Visualize results with graphs and confusion matrices. - Summarize findings and performance improvements. -Read and analyze relevant extra articles(if necessery) -Equaly separate task among the team -Discass with team given task and try to find common solution	End of Week 9

Week 10	**Reporting and Final Review** **&** **Create a final paper**	- Write the final report (results, conclusions, limitations). - Review and format the document according to submission guidelines. - Prepare the presentation or defense of the project.	End of Week 10

Research objective is to explore and identify the best method that will increase the quality of image before applying to CNN , AI detection of sick or healthy patient's by using their x-ray images. Our goal is to build pipline by using existing methods such as OpenCV, Numpy, Matplotlib , scikit-image that will increase the quality of image by scaling , normalizing , decrease the noise and so on .

Ethical Aspects in Preparing X-ray Images for AI Systems

While working on the topic **"Preparing high-quality X-ray images for AI-based diagnostic systems"**, I try to strictly follow the main ethical and responsible principles of using artificial intelligence in medicine. Ethics are especially important here because the accuracy of diagnoses, patient safety, data privacy, and public trust in medical AI all depend on it.

As Tianming Liu, Eliot Siegel, and Dinggang Shen mention, *"applications of AI-based intelligent and autonomous systems in radiology can increase the risk of systemic errors and raise complex ethical and societal issues."* This means that when preparing medical images, it's important to consider not only technical but also moral and social consequences of using AI.

In my work, I focus on the principles outlined by organizations like the **American College of Radiology (ACR)** and the **Radiological Society of North America (RSNA)**. According to their joint statement, *"ethical use of AI in radiology should promote well-being, minimize harm, and ensure that benefits and harms are distributed among stakeholders in a fair way."*

Based on these ideas, I follow several key rules when preparing X-ray images:

- **Confidentiality:** I always anonymize the images and make sure that patient privacy is protected.

- **Fairness:** I try to avoid any kind of bias in the data, for example, based on gender, age, or ethnicity.
- **Transparency:** I document all image processing steps so that the workflow is understandable and reproducible.
- **Responsibility:** I keep in mind that the quality of data directly affects how safely and accurately AI systems work.

As Danai Khemasuwan, Jeffrey S. Sorensen, and Henri G. Colt write, *"AI provides opportunities to improve the quality of care and accelerate the development of precision medicine. However, its limitations have led to the need for AI ethics and studying its impact on people and society."* This shows that technology alone is not enough — we also need to think about the ethical and social sides of AI in medicine.

One of the most important ethical principles I follow is **avoiding systemic errors and patient harm**. The authors warn that *"algorithms based on a small number of synthetic fictional cases... resulted in erroneous treatment recommendations."* For this reason, I make sure that all images are verified and of high quality, and I never include artificial or untested cases in my dataset.

Another key aspect is **fairness and eliminating bias**. It is mentioned that *"AI algorithms may be biased against patients of certain ethnicities or socioeconomic groups."* To prevent this, I prepare datasets that include people from different backgrounds and represent a realistic variety of patients.

I also pay special attention to **data security and confidentiality**, since *"the future of AI-related medical applications depends on how well safety, confidentiality and data security can be assured."* Therefore, I delete all identifiable information and use safe storage methods to protect data from leaks.

In addition, I agree with the idea that *"external validation using an independent dataset is critical before applying AI in real practice."* This step is necessary to make sure that models trained on my datasets are reliable and perform well in real-world conditions.

Finally, I completely support the statement that *"AI should respect human rights and freedoms, including dignity and privacy, and be designed for maximum transparency and dependability."* That's why in my work I follow the rule that accuracy and image quality should never come at the cost of patient rights or privacy.

To sum up, I believe that ethical principles in preparing X-ray images are not just formal requirements — they define how responsibly we use AI in medicine. I try to ensure confidentiality, fairness, transparency, and accountability at every stage of the process so that the systems built on this data truly serve people's health and safety.

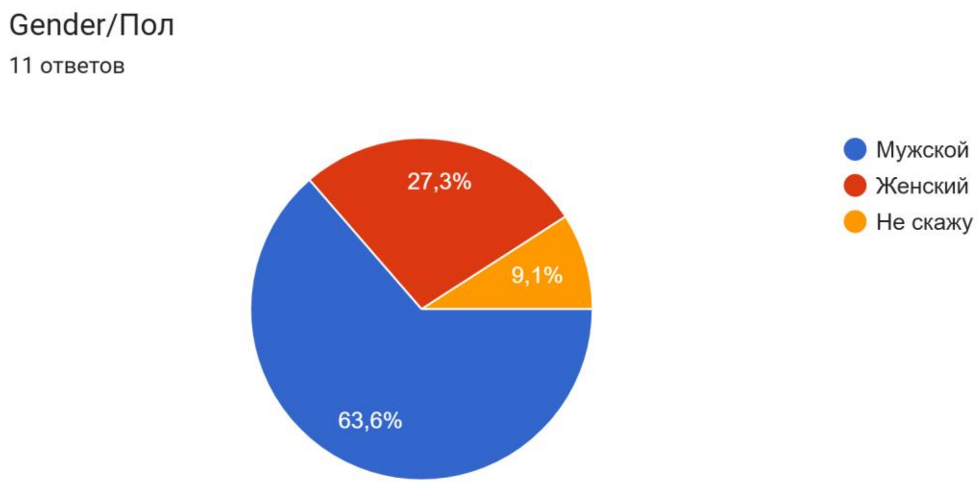
Questionnaire

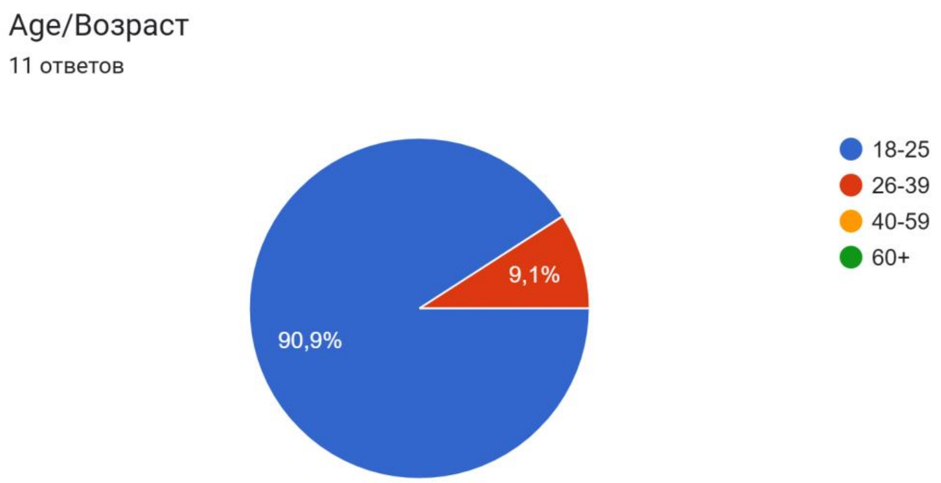
1. Questionnaire Objective and Design

A specialized survey was designed and administered to assess the perceptions, risks, and necessity of using artificial intelligence (AI) for enhancing the quality of X-ray images. The questionnaire's objective is to gather opinions from **all people** regarding the implementation of this technology into clinical practice.

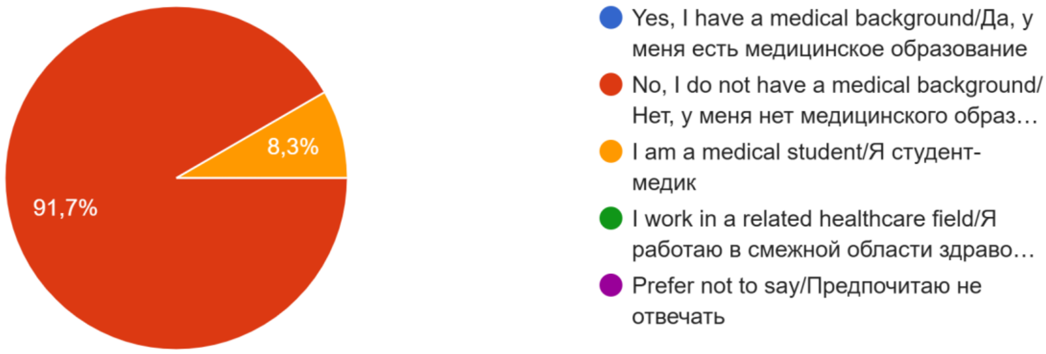
2. Demographic Data

The questionnaire included a section to collect anonymous demographic data. The purpose of this section was to understand the composition of the respondent sample and to analyze whether responses differed based on the participant's professional background and role.





Do you have a medical background?/ Есть ли у вас медицинское образование?
12 ответов

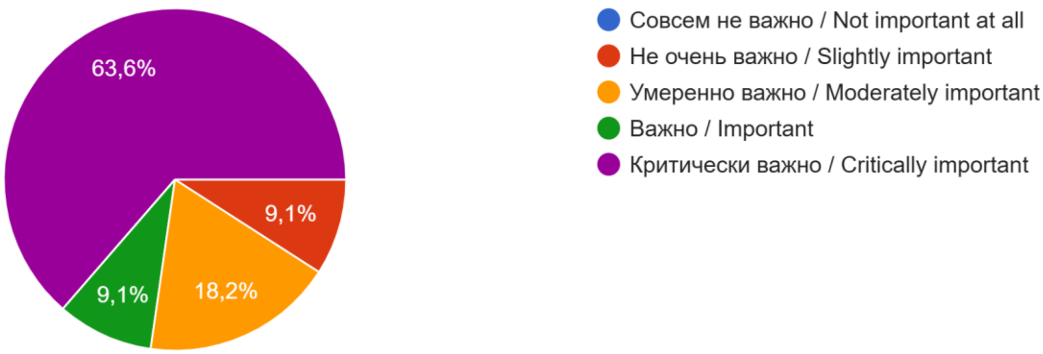


3. Questions

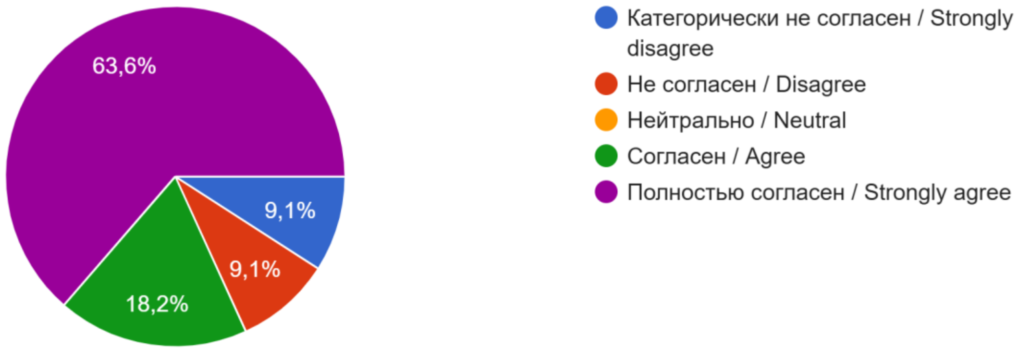
The questionnaire consists of 6 key questions designed to measure the following constructs:

1. **Problem Perception:** Assessing the perceived importance of image quality and the impact of poor quality on diagnostic errors.

1. Насколько важно, по вашему мнению, качество рентген-снимков для точности диагностики заболеваний? / How important do you ...uality of X-ray images for diagnostic accuracy?
11 ответов



3. Считаете ли вы, что низкое качество рентген-снимков может привести к ошибкам при диагностике? / Do you think that poor quality X-rays can lead to diagnostic errors?
11 ответов



2. **Necessity Assessment:** Determining the priority of developing AI enhancement tools compared to other solutions.

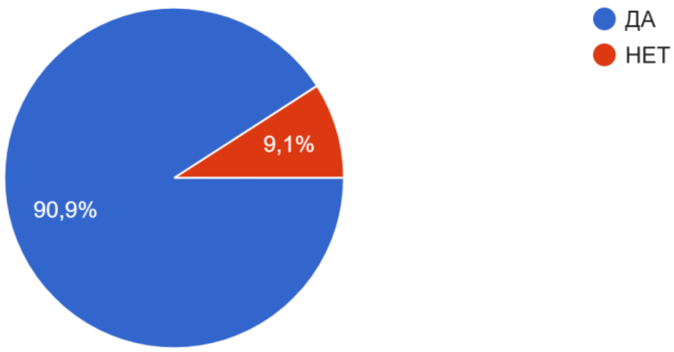
2. Насколько необходимой вы считаете разработку ИИ для улучшения качества снимков по сравнению с другими решениями / How necessary d... image quality is compared to other solutions?
11 ответов

3. **Risk and Trust:** Measuring the level of trust in AI regarding the potential creation of artifacts.

4. Насколько вы доверяете тому, что ИИ-модель (обученная улучшать снимки) не добавит на изображение детали, которых не было в исходных д...e that were not present in the original data?
11 ответов

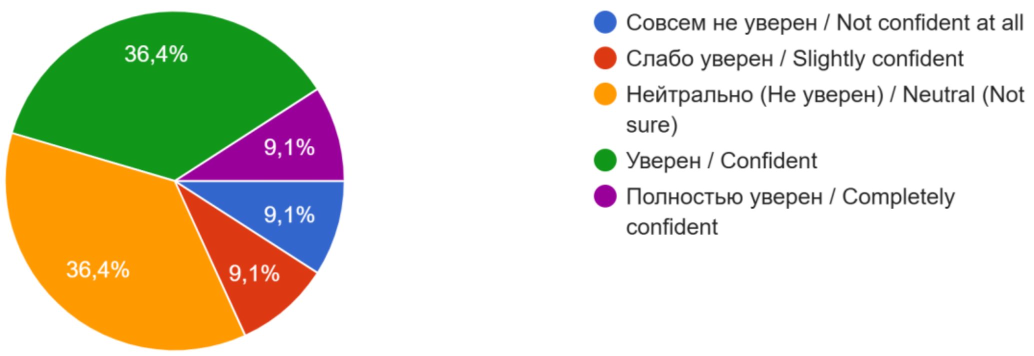
4. **Implementation Requirements:** Evaluating the necessity of retaining access to the original image for safety and transparency.

5. Должен ли врач всегда иметь доступ к оригинальному (до улучшения) снимку для сравнения с улучшенным? / Should the clinician ...ent) image to compare with the enhanced one?
11 ответов



5. **Efficacy Perception:** Gauging confidence in AI's ability to make unreadable images diagnostically usable.

6. Насколько вы уверены, что ИИ-улучшение может сделать нечитаемый (очень низкого качества) снимок диагностический пригодным? / ...(very low quality) image diagnostically usable?
11 ответов



4. Scaling and Data Collection

All items utilized a 5-point Likert scale to measure the degree of agreement, importance, necessity, or confidence. This scale was chosen to quantify subjective opinions. The response options were provided in both Russian and English to accommodate a broader audience.

3. Analysis Plan

The collected data will be analyzed using descriptive statistics for each question. This will allow for the identification of general trends and the level of consensus among respondents on the key aspects of using AI for image quality enhancement.

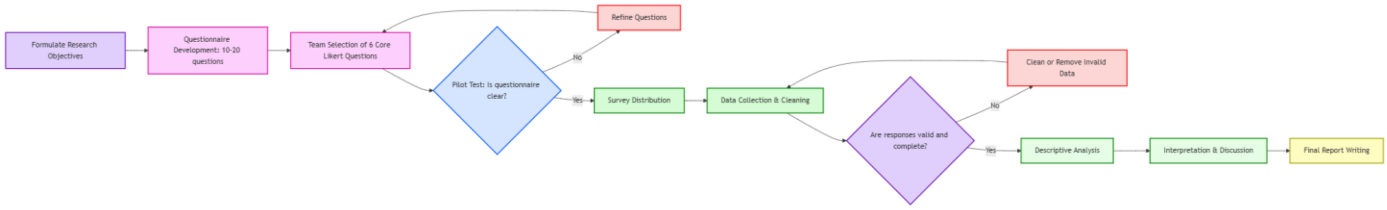


Figure 1. Survey Research Methodology Workflow

Interview Methodology

The questions were prepared in clear and simple language to encourage open discussion and to avoid technical ambiguity. Below are the main interview questions:

1. How often do you encounter low-quality chest X-ray images in your practice?
2. How does image quality affect the accuracy of diagnosis?
3. Which normalization methods do you consider to be the most effective for medical images?
4. How would you build a basic pipeline to improve the quality of X-ray images using tools like OpenCV, NumPy, Matplotlib, and scikit-image?
5. Do you think a standardized image preprocessing pipeline could help make diagnostic results more consistent?
6. How important is it to retain access to the original image after AI enhancement?
7. If you were creating your own preprocessing standard for X-ray, what steps would it definitely include?

The collected responses will be analyzed qualitatively to identify recurring themes and expert perspectives regarding data quality, diagnostic trust, and ethical considerations. The doctor's professional feedback was used to support the conclusions of this research and to validate the proposed approach to improving X-ray image quality for AI-based diagnostic systems.

Machine Learning Methodology

1.Dataset Identification and Description

This research uses two open-access chest X-ray databases - **COVID-19 Radiography Database** and **Chest X-ray Pneumonia Dataset** hosted on *Kaggle*. They were selected because of their complementary characteristics in terms of size, diversity, and data quality, which allow direct comparison of preprocessing performance across heterogeneous sources. The Dataset also gain popularity in site based on statistics.

COVID-19 Radiography Database:

Attribute	Details
Total Samples	33,920 chest X-ray (CXR) images
Classes	11,956 COVID-19, 11,263 Non-COVID infections (Viral or Bacterial Pneumonia), and 10,701 Normal
Image Format	Portable Network Graphics (PNG)
Image Size / Resolution	299*299 pixels.
Source	Aggregated from 43 medical publications, Kaggle datasets, Radiological Society of North America (RSNA) and etc.
Annotations	Radiologist-verified; each file labelled by disease type
Structure	/COVID/, /Lung_Opacity/, /Viral_Pneumonia/, /Normal/ sub-folders
License	Academic / Non-Commercial Use

Rationale for Selection:

This dataset provides a *large and diverse* testbed for developing normalization and noise-reduction pipelines.

The combination of four disease classes and heterogeneous acquisition sources enables the study to analyze how preprocessing affects both *data consistency* and *model robustness*.

Chest X-ray Pneumonia Dataset:

Attribute	Details
Total Samples	5,863 CXR images
Classes	Normal (1,583), Pneumonia (4,273)
Image Format	JPEG
Source	Guangzhou Women and Children's Medical Center, China
Annotations	Labelled and validated by certified pediatric radiologists
Structure	/train/, /test/, /val/ folders, each split into /NORMAL/ and /PNEUMONIA/
License	Public, for academic research
Image Quality	Consistent; minimal artefacts; proper field of view alignment

Rationale for Selection:
This is the control group dataset, due to its relatively clean and homogeneous structure, to which the results of the impact of preprocessing are compared.
This will allow performance differences to highlight, when processed through the same pipeline as the COVID-19 set, how cleaning and normalization improved the noisy real-world data.

Comparative Overview

Feature	COVID-19 Radiography	Chest X-ray Pneumonia
Images (total)	21,165	5,863
Classes	4	2
Format	PNG	JPEG
Acquisition Sources	Multi-hospital, international	Single hospital, China
Quality Variation	High (mixed)	Low (consistent)
Main Use in Study	Testing preprocessing on heterogeneous data	Baseline for clean data comparison

Preprocessing Approach

The preprocessing stage aims to prepare high-quality, standardized chest X-ray images that can be reliably used in AI-based diagnostic systems. Since image quality strongly affects the accuracy and performance of machine learning models, this phase focuses on reducing noise, improving contrast, and ensuring image consistency across the datasets.

Tools & libraries

Data preprocessing will be performed in Google Colaboratory (Colab). This choice is motivated by Colab's cost-free accessibility and its intuitive, browser-based environment, which is optimally suited for the efficient execution of Python-based data workflows and preliminary data manipulation tasks.

The computational tasks will rely on a focussed set of Python libraries, selected for their especial utility in image and numerical processing:

OpenCV - Open Computer Vision Library. Applied for basic image processing, such as resizing, filtering, and contrast stretching.

NumPy: It provides the realization of high-speed pixel-level mathematics and data normalization.

Matplotlib: Used here for the visualization of images and comparing visual data, and to demonstrate the result of some preprocessing steps.

scikit-image: Wrapped for advanced image quality enhancements and computation of quantitative evaluation metrics, such as the Structural Similarity Index (SSIM).

These libraries are chosen because they are lightweight, well-documented, and widely used in both academic and professional image analysis projects. They can be easily implemented by an individual researcher without requiring high computational power.

Preprocessing Steps

Each image from both datasets (COVID-19 Radiography Database and Chest X-ray Pneumonia Dataset) will be processed through the same simple and reproducible pipeline:

Step	Description	Purpose
1. Resize	Resize all images to a fixed resolution of 224×224 pixels.	Ensures uniform input size for later analysis.
2. Grayscale conversion and normalization	Convert all images to grayscale and scale pixel values to the [0, 1] range.	Standardizes intensity and simplifies model processing.
3. Noise reduction	Apply Gaussian blur filtering.	Minimizes random noise and artifacts.
4. Contrast enhancement	Use CLAHE (Contrast Limited Adaptive Histogram Equalization).	Improves visibility of lung features such as opacities or nodules.
5. Quality verification	Evaluate images using metrics such as brightness mean and SSIM.	Confirms improvement in image clarity and consistency.

Process Layout

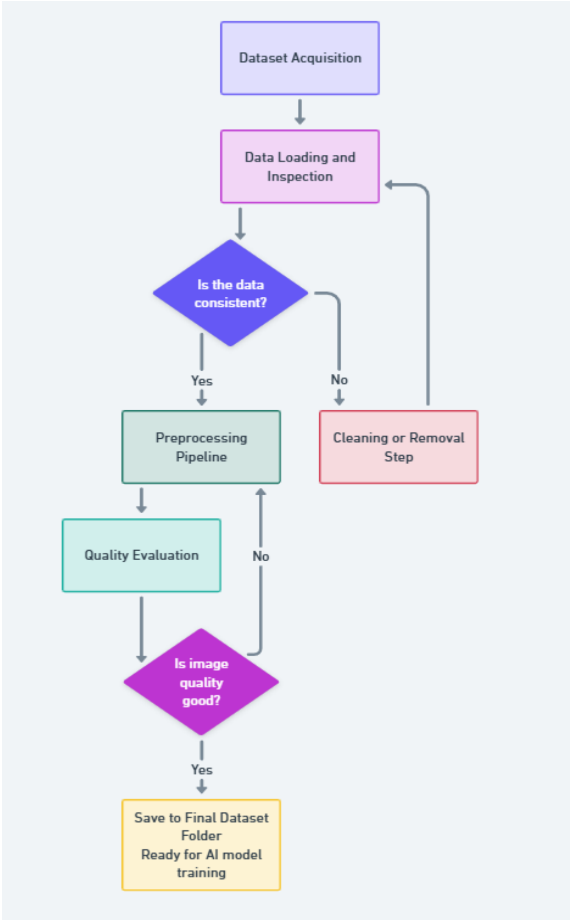


Figure 2. Data Processing and Quality Evaluation Workflow

The figure shows the main workflow for preparing high-quality chest X-ray images. The process starts with dataset acquisition and inspection in order to check consistency and remove corrupted samples. The clean data then enters a preprocessing pipeline, resizing, normalizing, noise reduction, and enhancing contrasts. Finally, images are tested in the quality evaluation stage using brightness and SSIM metrics. If quality is poor, images are subject to re-processing, while if it is good, they get saved into the final dataset, ready for AI training. This ensures that only clear and standardized images go into each AI analysis.

References

[1] T. Rahman, A. Khandakar, M. E. H. Chowdhury, Y. Qiblawey, A. Tahir, S. Kiranyaz *et al.*, "COVID-19 Radiography Database," *Kaggle Datasets*, 2021. [Online]. Available: https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database. [Accessed: 19-Oct-2025].

[2] P. Mooney, “Chest X-Ray Images (Pneumonia),” *Kaggle Datasets*, 2018. [Online]. Available: https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia. [Accessed: 19-Oct-2025].

[3] E. Elyan, P. Vuttipittayamongkol, P. Johnston, K. Martin, K. McPherson, C. F. Moreno-García, C. Jayne, and M. M. K. Sarker, “Computer vision and machine learning for medical image analysis: recent advances, challenges, and way forward,” *Artificial Intelligence Surgery*, vol. 2, pp. 24-45, 2022. https://www.oaepublish.com/articles/ais.2021.15

[4] J. Olveres, G. González, F. Torres, J. C. Moreno-Tagle, E. Carbajal-Degante, A. Valencia-Rodríguez, N. Méndez-Sánchez, and B. Escalante-Ramírez, “What is new in computer vision and artificial intelligence in medical image analysis applications,” *Quantitative Imaging in Medicine and Surgery*, vol. 11, no. 8, pp. 3830-3853, Aug. 2021. https://pmc.ncbi.nlm.nih.gov/articles/PMC8245941/

[5] T. Liu, E. Siegel, and D. Shen, “Deep learning and medical image analysis for COVID-19 diagnosis and prediction,” *Annual Review of Biomedical Engineering*, vol. 24, pp. 179-201, Jun. 2022. https://www.annualreviews.org/content/journals/10.1146/annurev-bioeng-110220-012203

[6] D. Khemasuwan, J. S. Sorensen, and H. G. Colt, “Artificial intelligence in pulmonary medicine: computer vision, predictive model and COVID-19,” *European Respiratory Review*, vol. 29, no. 157, art. 200181, 2020.https://publications.ersnet.org/content/errev/29/157/200181

Appendices

Appendix A— Tools and Environment Info

Tool	Purpose
Google Colab	Cloud-based Python environment
OpenCV	Image preprocessing
NumPy	Pixel normalization and analysis
scikit-image	SSIM computation
Matplotlib	Visualization

Appendix B— Link to our questionare:

https://forms.gle/o9zFzpEmFaWZzp1i6

Appendix C — Quality improvement techniques

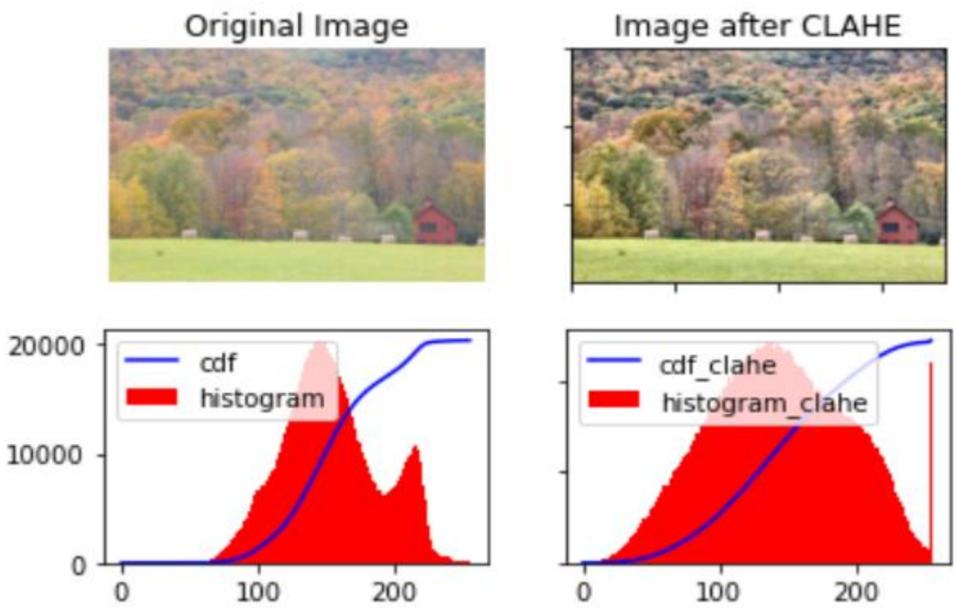


Figure 1. Contrast enhancement using CLAHE
The CLAHE algorithm increases local contrast, revealing finer lung structures and improving visibility of opacities.

Types of Contrast Enhancement Algorithms and Implementation in Python

Figure 2. Noise reduction using bilateral filter
Noise artifacts are reduced while preserving edges, resulting in a smoother and more consistent image.

[Bilateral filter - Wikipedia]

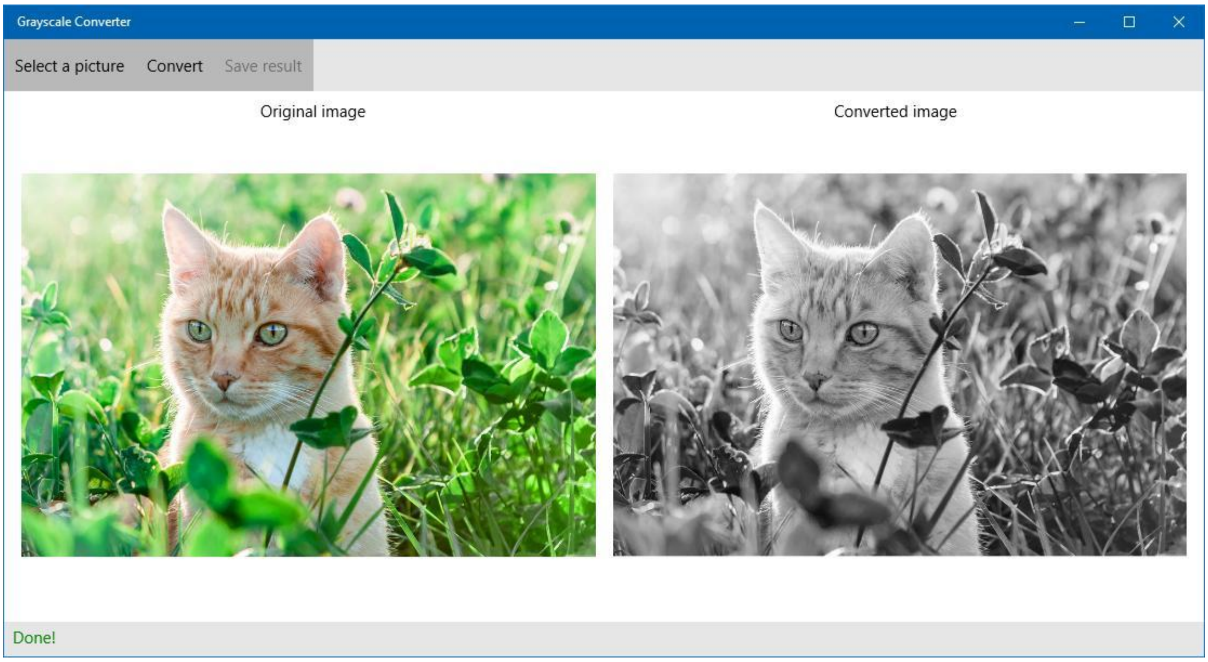


Figure 3. Grayscale conversion of chest X-ray
Conversion from RGB to grayscale simplifies the image by focusing on pixel intensity and structural details of the lungs.

[Why to use Grayscale Conversion during Image Processing?]