1. **Title: Improving radiological images quality through super resolution generative adversarial networks.**
2. **Research question**

Can the performance of current foundational models for image processing and quality enhancement be matched or improved by developing a model specifically tailored to radiological images?

1. **Abstract**

This research explores the enhancement of radiological image quality using a tailored generative adversarial network (GAN) model. Radiological images such as CT scans, MRIs, and X-rays often suffer from quality issues like low resolution, noise, and artifacts, which can impede accurate medical diagnoses and treatments. The study aims to develop a deep learning model specifically designed to address these deficiencies by leveraging GANs to improve image resolution and clarity. The proposed model will be trained on multimodal radiological images from public repositories, ensuring transparency in image transformations and interoperability with existing clinical workflows. By improving the visual quality of radiological images, this research seeks to enhance diagnostic accuracy, reduce the need for repeat studies, and optimize healthcare resources, ultimately promoting the adoption of AI technologies in clinical settings.

1. **Justification**

What problem is being addressed?

In current clinical environments, a significant proportion of radiological images—including CT scans, MRIs, and X-rays—exhibit quality deficiencies such as low resolution, noise, artifacts, or poor calibration. These defects affect medical interpretation, hinder accurate diagnoses, and may lead to inadequate treatments or the need to repeat studies (Boita et al., 2021). For instance, a retrospective review in the USA showed that 68% of breast MRI cases had at least one technical deficiency, with artifacts (74%) and low signal to noise ratio (mostly) being the most common (Ruiz-Flores et al., 2020)

Why does this problem exist?

The causes of poor quality in radiological images are diverse:

* Technological limitations of some acquisition equipment (Van et al., 2020).
* Patient conditions (movement, difficulty cooperating, presence of implants) (Sonni et al., 2018).
* Errors in image capture or equipment calibration (Mabotuwana et al., 2018).
* High demand for imaging services, necessitating rapid work without repeating defective captures.
* Lack of standardization in acquisition protocols.

Additionally, although artificial intelligence techniques have gained relevance in medicine, their application to improve image quality is still scarce in real-world settings, mainly due to the lack of interpretability of the results generated by these tools (Langlotz et al., 2019).

Why is it important to solve this problem?

The goal is to improve diagnostic efficiency and accuracy in the healthcare system, reducing errors, avoiding the repetition of studies, and decreasing the burden on medical personnel. Solving this problem will allow:

* Increased quality of images used in clinical diagnosis.
* Improved reliability of medical decisions.
* Optimized use of healthcare resources, reducing operational costs.
* Strengthened adoption of artificial intelligence technologies in real clinical settings.
* Promoted equitable and safe care, especially in populations with limited access to cutting-edge technology.

How is the problem proposed to be solved?

An artificial intelligence tool based on deep generative models, in this case, generative adversarial networks (GANs) (Yi et al., 2018), will be developed to enhance the visual quality of radiological images. The model will be trained with images from public repositories which have the following characteristics:

* Multimodality: It will work with different types of radiological images (X-rays, CT scans, MRIs).
* Explainable approach: Transparency in the transformations performed on the images will be sought.
* Interoperability: Designed to easily integrate into existing clinical workflows and systems.

The development of this tool will allow the recovery of images that would otherwise be discarded or considered of low diagnostic utility, offering a robust, scalable solution aligned with the real needs of the healthcare system (Chen et al., 2022).

1. **State of the art**

Multiple studies show how artificial intelligence tools and techniques have significantly accelerated in recent years, impacting various sectors and disciplines. This expansion is driven by advancements in machine learning, data availability, and computing power. AI has moved beyond prediction and summarization to generating content that resembles human work, becoming integrated into numerous societal sectors1. The field has shown a consistent upward trend, with significant growth in publications and applications (J. Liu et al., 2018).

Generative artificial intelligence (GenAI) is increasingly being recognized for its transformative potential in the field of medicine. By leveraging advanced machine learning techniques, GenAI can create new content and insights that enhance various aspects of healthcare, from diagnostics to education and patient care. Generative adversarial networks (GANs), a type of GenAI, are used in ophthalmology for image enhancement, disease identification, and synthetic data generation. These applications improve the accuracy and efficiency of ocular imaging techniques (Waisberg et al., 2024).

The integration of GenAI in healthcare raises significant privacy and security concerns. These include threats to protected health information and the need for strategies to mitigate these risks (Chen & Esmaeilzadeh, 2024), GenAI provides rapid access to information, improves diagnostic and prediction accuracy, and offers valuable insights into health conditions. It is particularly beneficial when healthcare systems adapt culturally to integrate this technology (Jindal et al., 2024)

Computer vision has become a pivotal tool in medical imaging, offering significant advancements in diagnostic assistance. By leveraging artificial intelligence and machine learning, computer vision systems can enhance the accuracy and efficiency of medical diagnoses across various imaging modalities. Computer vision techniques have been applied to facial image analysis for medical diagnostics, enabling the detection of over 30 medical conditions through the identification of facial abnormalities. However, the clinical adoption of these methods is limited due to concerns about reliability and the need for further validation and applicability in real-world settings (Thevenot et al., 2018). Computer vision systems, particularly those using deep learning, have shown promise in segmenting and classifying medical images such as CT and MRI scans. For instance, convolutional neural networks (CNNs) have been effectively used for lung segmentation in CT images, achieving high accuracy and efficiency (Hu et al., 2020).

Despite the advancements, integrating computer vision systems into clinical practice remains challenging. Issues such as the need for large, annotated datasets for training, and the adaptation of algorithms to diverse clinical environments, are significant hurdles (Schmidt et al., 2023). Moreover, ensuring the systems' reliability and gaining clinical validation are critical for widespread adoption (Olveres et al., 2021).

Super-resolution (SR) is a technique used to enhance the resolution of images or videos beyond the limitations of the original imaging system. Recent advances in super-resolution have been driven by developments in deep learning and other computational methods (Lepcha et al., 2022). The introduction of deep learning has significantly advanced SR capabilities. Techniques such as convolutional neural networks (CNNs) and transformer-based models have improved the quality and efficiency of SR processes, allowing for better handling of complex images and videos (Chauhan et al., 2020).

1. **Theoretical framework**

Diagnostic imaging (Morris & Perkins, 2012) refers to medical techniques that allow visualization of the interior of the human body non-invasively, using technologies such as X-rays (Ou et al., 2021), computed tomography (CT) (Lu et al., 1977), magnetic resonance imaging (MRI) (Oppelt & Grandke, 1993), ultrasound, and nuclear medicine. These tools are fundamental for the detection, diagnosis, and monitoring of a wide variety of diseases, from bone fractures and tumors to cardiovascular and neurological conditions. Globally, diagnostic imaging plays a crucial role in improving healthcare by facilitating more accurate diagnoses and more effective treatments, thus reducing mortality and improving the quality of life for millions of people. Additionally, its continuous development drives technological innovation and the advancement of personalized medicine.

Radiation is used in several diagnostic imaging techniques because it has the ability to penetrate the human body and generate contrasts between different tissues, allowing detailed visualization of internal structures (Dendy & Heaton, 1999). For example, in X-rays and computed tomography (CT) scans, X-rays—a form of electromagnetic radiation—pass through the body and are absorbed to varying degrees by bones, organs, and tissues. These absorption levels translate into grayscale images that reveal internal anatomy. This facilitates the diagnosis and identification of various pathologies or anomalies in patients. This is done through various techniques, some of which use radiation and some do not. The most common are as follows:

* Radiography (X-rays): Uses electromagnetic radiation to penetrate the body. Dense tissues like bones absorb more X-rays and appear white, while soft tissues allow more radiation to pass through and appear darker in the image. It is quick and useful for evaluating bones, lungs, and thoracic structures.
* Computed Tomography (CT): Uses X-rays combined with computational processing to generate cross-sectional images (slices) of the body. It offers more detail than a simple X-ray and is ideal for studying internal organs, trauma, and tumors.
* Magnetic Resonance Imaging (MRI): Uses powerful magnetic fields (tesla) and radio waves to align the body's protons and then record their behavior as they return to their normal state. It does not use ionizing radiation and provides very detailed images of soft tissues such as the brain, muscles, joints, and spinal cord.
* Ultrasound (sonography): Emits high-frequency sound waves that bounce off body tissues. The echo of these waves is converted into real-time images. It is safe, portable, and widely used in obstetrics, abdominal, and cardiac studies.
* Nuclear Medicine (scintigraphy, PET): Introduces small amounts of radioactive material (radiopharmaceuticals) into the body. These emit radiation that is captured by special cameras to show the functioning of organs and tissues, rather than their structure. It is key in oncology, cardiology, and neurology.

The procedure for taking diagnostic images varies depending on the technique used, but generally follows common steps that ensure patient safety and image quality:

* Patient preparation: Depending on the type of image, the patient may need to follow specific instructions, such as fasting, not wearing jewelry, or dressing in a medical gown. In some studies, a contrast medium (oral, intravenous, or rectal) is used to improve the visualization of certain internal structures.
* Positioning: The patient is carefully placed in the appropriate position on the table or inside the imaging equipment (for example, in the MRI tube or under the CT scanner arch), depending on what needs to be observed.
* Image acquisition: The technical staff operates the equipment from a control room. During image acquisition, the patient must remain still (and sometimes hold their breath for a few seconds) to avoid blurry images. The duration of the study can range from a few minutes (simple X-rays) to more than an hour (complex MRI).
* Processing and analysis: The obtained images are digitally processed and sent to a storage system (PACS), where a specialized radiologist analyzes and interprets them.
* Diagnostic report: The radiologist issues a detailed medical report that is sent to the treating physician, who uses it to guide treatment or make clinical decisions.

Computer vision is a branch of artificial intelligence that allows machines to interpret and understand images and videos of the real world similarly to the human eye (Pulli et al., 2012). It uses advanced algorithms to detect patterns, extract features, recognize objects, and make decisions based on visual information.

This technology is applied in numerous fields such as industry, security, autonomous transportation, and medicine, among others. Its main strength is the ability to process large volumes of visual data quickly and accurately, making it a key tool for automating complex tasks that previously required human intervention. Computer vision techniques applied in clinical settings enable automated and precise analysis of medical images, facilitating the detection, segmentation, and classification of anatomical structures and possible pathologies (Esteva et al., 2021)

These techniques use digital processing algorithms and artificial intelligence, such as convolutional neural networks, to interpret large volumes of visual data quickly and accurately.

Their relevance lies in their ability to support medical diagnosis, reduce human error, streamline clinical workflow, and enable early diagnoses, even in regions with a shortage of specialists (Teslenko & Smelyakov, 2024). Thanks to computer vision, it is possible to develop systems for diagnosis support, disease monitoring, and treatment personalization, marking a significant advance towards more efficient, precise, and accessible medicine.

Among the different computer vision techniques, a recent advancement is super-resolution, which is an image processing technique aimed at increasing the spatial resolution of an image, i.e., improving its level of detail and sharpness from a low-resolution image (Capel & Zisserman, 2003). This is achieved through algorithms that estimate and reconstruct missing information based on learned patterns or statistical inferences. Traditional methods such as interpolation exist, but more modern approaches use deep neural networks that learn to generate high-quality versions from large volumes of data.

Neural networks are computational models inspired by the functioning of the human brain, composed of layers of nodes or "neurons" that process information through weighted connections (Prieto et al., 2016). These networks can learn complex patterns from data through a training process that adjusts the weights of the connections to minimize prediction error.

Deep learning (LeCun, 2019) is a subfield of machine learning (Janiesch et al., 2021) that uses neural networks with many hidden layers—hence the term "deep"—to model highly nonlinear and abstract relationships. Thanks to its ability to learn hierarchical representations of data, deep learning has revolutionized fields such as computer vision, natural language processing, and medical image analysis. Its success is partly due to access to large amounts of data and modern computational power, which have enabled the training of deeper and more accurate models than ever before.

SRGAN (Super-Resolution Generative Adversarial Network) is a deep neural network model used to enhance image resolution through a technique called super-resolution. SRGAN is based on the architecture of generative adversarial networks (GAN)(Creswell et al., 2017), which consists of two main components: a generator and a discriminator.

* Generator: Its task is to create high-resolution images from low-resolution images. It uses a deep neural network that learns to generate sharper and more textured details that resemble those present in high-quality images.
* Discriminator: Its task is to distinguish between the images generated by the model (high-resolution) and real high-resolution images. During training, the discriminator tries to identify whether an image is real or fake, while the generator tries to deceive it by producing increasingly realistic images.

The training process is adversarial, meaning that both components compete against each other: the generator tries to improve its images to deceive the discriminator, while the discriminator tries to improve its task of distinguishing generated images from real ones (Wang et al., 2017). This approach gradually improves the quality of the generated images.

A notable feature of SRGAN is the use of content loss and adversarial loss. Content loss ensures that the generated images are visually similar to the original high-resolution image, while adversarial loss allows the generator to produce more realistic images, thanks to the feedback provided by the discriminator. Thanks to these losses, SRGAN is capable of producing high-resolution images with fine details and realistic textures, making it very useful in applications such as medical image enhancement, photography, and restoration of old images.

Of course, integrating deep learning models is not easy, as they usually have this “black box” nature, where due to the amount and complexity of the operations between nodes and layers, makes it impossible to be understandable or interpretable for any human. From this flair comes the necessity to explore options for interpretable deep learning (Li et al., 2021) , through techniques that allow visualizing, explaining, or justifying the decisions of the model, trust and adoption of these technologies in sensitive areas such as medicine, banking, or justice are facilitated. Some of these techniques include heatmaps, visual attention, feature importance analysis, and methods like LIME or SHAP. The main goal is for human experts, such as doctors or engineers, to understand why a model reached a certain conclusion, detect possible errors or biases, and validate whether the decisions are aligned with professional knowledge and ethics (Li et al., 2021).

In order to accomplish this task, various metrics are being implemented to monitor not only the model performance, but also, see through its complexity and understanding the reasons behind the decision making:

LIME (Local Interpretable Model-Agnostic Explanations) is an interpretability technique used to explain the predictions of complex models (such as neural networks) in a way that is understandable to humans. It does this locally, meaning it explains the prediction of a single data instance rather than trying to understand the entire model. LIME works by generating a set of synthetic data around an input instance, applying small perturbations to the data (e.g., modifying features of the instance), and observing how the model's prediction changes. It then fits an interpretable model (such as linear regression) to these synthetic data and uses this interpretable model to provide an explanation of the original prediction (Rahnama et al., 2024).

SHAP (Shapley Additive Explanations) is based on the theory of Shapley values, a concept from game theory that assigns a "fair contribution" to each feature in a model. The central idea of SHAP is to calculate how the model's prediction changes when a particular feature is included or excluded. By doing this for all features, SHAP provides an accurate measure of the importance of each feature in the prediction of an instance (Nohara et al., 2021).

Grad-CAM (Gradient-weighted Class Activation Mapping) is a technique used to visualize the areas of an image that contribute most to the prediction of a convolutional neural network (CNN). Grad-CAM generates a visual activation map based on the gradients of the class of interest with respect to the activations of the last convolutional layer of the network. This allows identifying the regions of the image that most influence the model's final decision (Selvaraju et al., 2016).

Cycle-GAN Consistency refers to a key property of this model: cycle consistency, which ensures that after transforming an image from one domain to another, reversing the process accurately recovers the original image. Cycle consistency is important because it ensures that the generator in the Cycle-GAN does not lose important information during the transformation and can revert the image back to its original state without significant errors (H. Liu et al., 2020).

**Objectives**

General

Develop a generative artificial intelligence model specifically designed to enhance the quality of radiological images, surpassing the performance of current foundational models in image processing.

Specifics

1. Implement deep learning models within a generative adversarial network (GAN) training framework to enhance image quality.
2. Design a tool to monitor and evaluate the model's training performance using various relevant metrics.
3. Validate the results by consulting radiology specialists to assess the improvement in image quality.
4. **Methodology**

Database Structuring

Given that the information related to the images quality used in this study is not structured, the images will be automatically classified as Low-resolution (LR) or High-resolution (HR) to be fed into the model.

The data will be organized and stored in a structured manner to facilitate access during the training, validation, and testing stages. The database structure will be designed following these guidelines:

Labelling and Quality Control: Images will be accompanied by a CSV file containing information about their quality and if they are restored or not.

Folder Structure: Images will be organized according to the following scheme: Train/, Validation/, and Test/, with a 70%-15%-15% distribution respectively. Each folder will contain subdirectories separated for original and restored images.

Data Analysis and Processing Images will be imported from various public repositories. During preprocessing, quality control will be applied using digital processing techniques, utilizing the OpenCV framework and other specialized libraries. This process will include:

* Intensity normalization.
* Resizing to a standard resolution
* Orientation detection and correction.
* Cropping unnecessary edges.
* Removal of redundant or corrupted images.
* Original and restored images will subsequently be aligned to allow fair comparisons in quality and perception evaluations.

Models to be Used

Generative Adversarial Networks (GAN): This architecture will be used for quality enhancement and restoration of old or low-quality radiological images. The generator network will learn to produce higher resolution images, while the discriminator network will evaluate their authenticity.

VGG19: Used as a perceptual network to calculate perceptual loss during GAN model training. VGG19 will provide a measure of visual similarity between generated and real images at deeper semantic levels.

Evaluation Metrics

Various quantitative and qualitative metrics will be used to evaluate the model's performance:

* Adversarial Loss: Evaluates how effectively the generator network deceives the discriminator.
* Perceptual Loss: Measures the semantic difference between the original and restored image using intermediate layers of VGG19.
* LIME (Local Interpretable Model-agnostic Explanations): An interpretability technique that provides local explanations about which regions of the image most influence the model's decision.
* PSNR (Peak Signal-to-Noise Ratio): A traditional quantitative metric that measures the reconstruction fidelity between two images.
* Interpretability Metrics To improve model transparency and facilitate clinical adoption, the following interpretability techniques will be used:
* Grad-CAM: This technique identifies the regions of an image that were most relevant in the model's decision-making. It will be applied to both the discriminator and auxiliary classifier networks, if any.
* Cycle-GAN Consistency: Ensures that generated images can be transformed back to their original state without significant loss of information, thus preserving relevant anatomical structures.

Evaluating models

After concluding our training, we need to compare our resulting enhanced images with the ones produced by commercial foundational models. This will require assistance from radiologist experts in order to have concluding results.

1. **Expected results**

In the last months, most updates for foundational models have been solely focused on image processing and generating. It is evident that those models nowadays can take images and improve them significantly. However, in medical institutions the number of images that are produced every day is quite significant, so using API to these models or implementing AI agents with these commercial models can be very expensive. Our goal is to determine if with a large number of public images, we can create a specialized model that can be equal or better than the commercial models when improving radiological images quality. As results of this project, we expect to get:

* Model: A robust model that actually improves radiological images quality
* Comparative table between our models and the commercial ones

1. **References**

Boita, J., Van Engen, R., Mackenzie, A., Tingberg, A., Bosmans, H., Bolejko, A., Zackrisson, S., Wallis, M., Ikeda, D., Van Ongeval, C., Pijnappel, R., Broeders, M., Sechopoulos, I., De Bruin Andersson Behmer Taylor Kilb, F. J. D., Jansen, F., Duijm, L., De Bruin, H., Andersson, I., Behmer, C., … Pal, S. (2021). How does image quality affect radiologists’ perceived ability for image interpretation and lesion detection in digital mammography? *European Radiology*, *31*, 5335–5343. <https://doi.org/10.1007/s00330-020-07679-8>

Capel, D., & Zisserman, A. (2003). Computer vision applied to super resolution. *IEEE Signal Process. Mag.*, *20*, 75–86. <https://doi.org/10.1109/MSP.2003.1203211>

Chauhan, K., Patel, H., Dave, R., Bhatia, J., & Kumhar, M. (2020). *Advances in Single Image Super-Resolution: A Deep Learning Perspective*. 443–455. <https://doi.org/10.1007/978-981-15-3369-3_34>

Chen, Y., & Esmaeilzadeh, P. (2024). Generative AI in Medical Practice: In-Depth Exploration of Privacy and Security Challenges. *Journal of Medical Internet Research*, *26*. <https://doi.org/10.2196/53008>

Chen, Y., Yang, X.-H., Wei, Z., Heidari, A. A., Zheng, N., Li, Z., Chen, H., Hu, H., Zhou, Q., & Guan, Q. (2022). Generative Adversarial Networks in Medical Image augmentation: A review. *Computers in Biology and Medicine*, *144*, 105382. <https://doi.org/10.1016/j.compbiomed.2022.105382>

Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., & Bharath, A. (2017). Generative Adversarial Networks: An Overview. *IEEE Signal Processing Magazine*, *35*, 53–65. <https://doi.org/10.1109/MSP.2017.2765202>

Dendy, P., & Heaton, B. (1999). Physics for Diagnostic Radiology. *Medical Physics*, *28*, 404. <https://doi.org/10.1118/1.1350686>

Esteva, A., Chou, K., Yeung, S., Naik, N., Madani, A., Mottaghi, A., Liu, Y., Topol, E., Dean, J., & Socher, R. (2021). Deep learning-enabled medical computer vision. *NPJ Digital Medicine*, *4*. <https://doi.org/10.1038/s41746-020-00376-2>

Hu, Q., Souza, L., Holanda, G., Alves, S., Silva, F., Han, T., & Filho, P. (2020). An effective approach for CT lung segmentation using mask region-based convolutional neural networks. *Artificial Intelligence in Medicine*, *103*, 101792. <https://doi.org/10.1016/j.artmed.2020.101792>

Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, *31*, 685–695. <https://doi.org/10.1007/s12525-021-00475-2>

Jindal, J., Lungren, M., & Shah, N. (2024). Ensuring useful adoption of generative artificial intelligence in healthcare. *Journal of the American Medical Informatics Association : JAMIA*. <https://doi.org/10.1093/jamia/ocae043>

Langlotz, C., Allen, B., Erickson, B., Kalpathy-Cramer, J., Bigelow, K., Cook, T., Flanders, A., Lungren, M., Mendelson, D., Rudie, J., Wang, G., & Kandarpa, K. (2019). A Roadmap for Foundational Research on Artificial Intelligence in Medical Imaging: From the 2018 NIH/RSNA/ACR/The Academy Workshop. *Radiology*, *291 3*, 781–791. <https://doi.org/10.1148/radiol.2019190613>

LeCun, Y. (2019). 1.1 Deep Learning Hardware: Past, Present, and Future. *2019 IEEE International Solid- State Circuits Conference - (ISSCC)*, 12–19. <https://doi.org/10.1109/ISSCC.2019.8662396>

Lepcha, D., Goyal, B., Dogra, A., & Goyal, V. (2022). Image super-resolution: A comprehensive review, recent trends, challenges and applications. *Inf. Fusion*, *91*, 230–260. <https://doi.org/10.1016/j.inffus.2022.10.007>

Li, X., Xiong, H., Li, X., Wu, X., Zhang, X., Liu, J., Bian, J., & Dou, D. (2021). Interpretable deep learning: interpretation, interpretability, trustworthiness, and beyond. *Knowledge and Information Systems*, *64*, 3197–3234. <https://doi.org/10.1007/s10115-022-01756-8>

Liu, H., Liu, J., Hou, S., Tao, T., & Han, J. (2020). Perception consistency ultrasound image super-resolution via self-supervised CycleGAN. *Neural Computing and Applications*, *35*, 12331–12341. <https://doi.org/10.1007/s00521-020-05687-9>

Liu, J., Kong, X., Xia, F., Bai, X., Wang, L., Qing, Q., & Lee, I. (2018). Artificial Intelligence in the 21st Century. *IEEE Access*, *6*, 34403–34421. <https://doi.org/10.1109/ACCESS.2018.2819688>

Lu, Y. I. S. A., Cl, Metu, F., & Metu, A. (1977). COMPUTERIZED TOMOGRAPHY. *Medical Journal of Australia*, *2*. <https://doi.org/10.1007/3-540-29832-0_397>

Mabotuwana, T., Bhandarkar, V., Hall, C., & Gunn, M. (2018). Detecting Technical Image Quality in Radiology Reports. *AMIA ... Annual Symposium Proceedings. AMIA Symposium*, *2018*, 780–788. <https://consensus.app/papers/detecting-technical-image-quality-in-radiology-reports-mabotuwana-bhandarkar/1fea73aa988354b28a0be5eae5e194c9/>

Morris, P., & Perkins, A. (2012). Diagnostic imaging. *The Lancet*, *379*, 1525–1533. <https://doi.org/10.1016/S0140-6736(12)60429-2>

Nohara, Y., Matsumoto, K., Soejima, H., & Nakashima, N. (2021). Explanation of Machine Learning Models Using Shapley Additive Explanation and Application for Real Data in Hospital. *Computer Methods and Programs in Biomedicine*, *214*, 106584. <https://doi.org/10.1016/j.cmpb.2021.106584>

Olveres, J., González, G., Torres, F., Moreno-Tagle, J. C., Carbajal-Degante, E., Valencia-Rodríguez, A., Méndez-Sánchez, N., & Escalante-Ramírez, B. (2021). What is new in computer vision and artificial intelligence in medical image analysis applications. *Quantitative Imaging in Medicine and Surgery*, *11 8*, 3830–3853. <https://doi.org/10.21037/QIMS-20-1151>

Oppelt, A., & Grandke, T. (1993). Magnetic resonance imaging. *Superconductor Science and Technology*, *6*, 381–395. <https://doi.org/10.1088/0953-2048/6/6/001>

Ou, X., Chen, X., Xu, X., Xie, L., Chen, X., Hong, Z., Bai, H., Liu, X., Chen, Q., Li, L., & Yang, H. (2021). Recent Development in X-Ray Imaging Technology: Future and Challenges. *Research*, *2021*. <https://doi.org/10.34133/2021/9892152>

Prieto, A., Prieto, B., Ortigosa, E., Ros, E., Pelayo, F., Ortega, J., & Rojas, I. (2016). Neural networks: An overview of early research, current frameworks and new challenges. *Neurocomputing*, *214*, 242–268. <https://doi.org/10.1016/j.neucom.2016.06.014>

Pulli, K., Baksheev, A., Kornyakov, K., & Eruhimov, V. (2012). Realtime Computer Vision with OpenCV. *Queue*, *10*, 40–56. <https://doi.org/10.1145/2181796.2206309>

Rahnama, A. H. A., Alfaro, L. G., Wang, Z., & Movin, M. (2024). *Local Interpretable Model-Agnostic Explanations for Neural Ranking Models*. 151–159. <https://doi.org/10.3384/ecp208017>

Ruiz-Flores, L., Whitman, G., Le‐Petross, H., Hess, K., & Parikh, J. (2020). Variation in Technical Quality of Breast MRI. *Academic Radiology*. <https://doi.org/10.1016/j.acra.2019.07.002>

Schmidt, A., Mohareri, O., DiMaio, S., Yip, M., & Salcudean, S. (2023). Tracking and Mapping in Medical Computer Vision: A Review. *Medical Image Analysis*, *94*, 103131. <https://doi.org/10.1016/j.media.2024.103131>

Selvaraju, R., Das, A., Vedantam, R., Cogswell, M., Parikh, D., & Batra, D. (2016). Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *International Journal of Computer Vision*, *128*, 336–359. <https://doi.org/10.1007/s11263-019-01228-7>

Sonni, I., Baratto, L., Park, S., Hatami, N., Srinivas, S., Davidzon, G., Gambhir, S., & Iagaru, A. (2018). Initial experience with a SiPM-based PET/CT scanner: influence of acquisition time on image quality. *EJNMMI Physics*, *5*. <https://doi.org/10.1186/s40658-018-0207-x>

Teslenko, D., & Smelyakov, K. (2024). ROLE AND EVOLUTION OF COMPUTER VISION IN MEDICINE. *Grail of Science*. <https://doi.org/10.36074/grail-of-science.15.03.2024.030>

Thevenot, J., López, M. B., & Hadid, A. (2018). A Survey on Computer Vision for Assistive Medical Diagnosis From Faces. *IEEE Journal of Biomedical and Health Informatics*, *22*, 1497–1511. <https://doi.org/10.1109/JBHI.2017.2754861>

Van, A. S. X. N. M. A. R. A.-B. S. G. deSouza S., Alberich-Bayarri, Á., Sourbron, S., Golay, X., deSouza, N., Smits, M., Van Der Lugt, A., & Boellard, R. (2020). ESR Statement on the Validation of Imaging Biomarkers. *Insights into Imaging*, *11*. <https://doi.org/10.1186/s13244-020-00872-9>

Waisberg, E., Ong, J., Kamran, S., Masalkhi, M., Paladugu, P., Zaman, N., Lee, A., & Tavakkoli, A. (2024). Generative artificial intelligence in ophthalmology. *Survey of Ophthalmology*. <https://doi.org/10.1016/j.survophthal.2024.04.009>

Wang, K., Gou, C., Duan, Y., Lin, Y., Zheng, X., & Wang, F.-Y. (2017). Generative adversarial networks: introduction and outlook. *IEEE/CAA Journal of Automatica Sinica*, *4*, 588–598. <https://doi.org/10.1109/JAS.2017.7510583>

Yi, X., Walia, E., & Babyn, P. (2018). Generative Adversarial Network in Medical Imaging: A Review. *Medical Image Analysis*, *58*, 101552. <https://doi.org/10.1016/j.media.2019.101552>