

Review



Artificial Intelligence in Chest Radiography—A Comparative Review of Human and Veterinary Medicine

Andrea Rubini ^{1,*}, Roberto Di Via ², Vito Paolo Pastore ², Francesca Del Signore ^{1,*}, Martina Rosto ¹, Andrea De Bonis ¹, Francesca Odone ² and Massimo Vignoli ¹

- ¹ Department of Veterinary Medicine, University of Teramo, Piano d'Accio, 64100 Teramo, Italy; mrosto@unite.it (M.R.); adebonis@unite.it (A.D.B.); mvignoli@unite.it (M.V.)
- ² MaLGa, DIBRIS, University of Genoa, Via Dodecaneso 35, 16146 Genoa, Italy;

roberto.divia@edu.unige.it (R.D.V.); vito.paolo.pastore@unige.it (V.P.P.); francesca.odone@unige.it (F.O.) * Correspondence: arubini@unite.it (A.R.); fdelsignore@unite.it (F.D.S.)

Simple Summary: Artificial intelligence (AI) could enhance the field of radiology in both human and veterinary medicine by making diagnoses faster and more accurate. In human healthcare, AI assists in detecting diseases such as pneumonia and COVID-19, supporting physicians in pattern recognition and outcome prediction. However, human oversight remains essential due to data limitations and ethical concerns. In veterinary medicine, the use of AI is still limited due to several factors, including the lack of large databases, anatomical differences between animal breeds, and limited research in this field. Focusing on species with less anatomical variability, such as cats, and encouraging interdisciplinary collaboration could foster its development. Despite its potential, the radiologist's expertise remains crucial. In this context, AI can be seen as a valuable support tool in the daily practice of radiology.

Abstract: The integration of artificial intelligence (AI) into chest radiography (CXR) has greatly impacted both human and veterinary medicine, enhancing diagnostic speed, accuracy, and efficiency. In human medicine, AI has been extensively studied, improving the identification of thoracic abnormalities, diagnostic precision in emergencies, and the classification of complex conditions such as tuberculosis, pneumonia, and COVID-19. Deep learning-based models assist radiologists by detecting patterns, generating probability maps, and predicting outcomes like heart failure. However, AI is still supplementary to clinical expertise due to challenges such as data limitations, algorithmic biases, and the need for extensive validation. Ethical concerns and regulatory constraints also hinder full implementation. In veterinary medicine, AI is still in its early stages and is rarely used; however, it has the potential to become a valuable tool for supporting radiologists in the future. However, challenges include smaller datasets, breed variability, and limited research. Addressing these through focused research on species with less phenotypic variability (like cats) and cross-sector collaborations could advance AI in veterinary medicine. Both fields demonstrate AI's potential to enhance diagnostics but emphasize the ongoing need for human expertise in clinical decision making. Differences in anatomy structure between the two fields must be considered for effective AI adaptation.

Keywords: artificial intelligence; deep learning; chest radiography



Academic Editors: Vitor Filipe, Lio Gonçalves and Mário Ginja

Received: 6 March 2025 Revised: 23 April 2025 Accepted: 24 April 2025 Published: 25 April 2025

Citation: Rubini, A.; Di Via, R.; Pastore, V.P.; Del Signore, F.; Rosto, M.; De Bonis, A.; Odone, F.; Vignoli, M. Artificial Intelligence in Chest Radiography—A Comparative Review of Human and Veterinary Medicine. *Vet. Sci.* **2025**, *12*, 404. https://doi.org/10.3390/ vetsci12050404

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/ licenses/by/4.0/).

1. Introduction

The term "artificial intelligence" (AI) refers to the use of computer systems designed to solve specific problems by simulating human reasoning [1]. A key trait of AI is its capacity to adjust solutions dynamically in response to evolving circumstances, mirroring human adaptability. While AI is designed to mimic specific cognitive functions, it often surpasses human capabilities in tasks such as processing vast amounts of data, recognizing complex patterns, and performing high-speed analyses [2]. AI operates through computational models and algorithms: structured sets of mathematical rules and coded instructions that transform input data into meaningful outputs, enabling problem-solving across various domains. Machine learning (ML) is a core AI methodology that enables computers to learn from examples rather than through explicit programming. By analyzing large datasets, these systems can identify patterns, generate predictions, and support decision-making processes [3]. Building upon these foundations, deep learning (DL) represents a significant advancement in AI research and now forms the basis for most recent AI innovations in the field. This methodology employs Artificial Neural Networks (ANNs) inspired by biological neural systems, processing information through a hierarchical structure comprising an input layer, multiple hidden layers, and an output layer. Through its distinctive incorporation of numerous hidden processing layers, deep learning efficiently analyzes complex, multidimensional veterinary data that traditional methods struggle to interpret. These capabilities have enabled sophisticated applications across veterinary medicine, including advanced diagnostic imaging analysis and complex clinical parameter interpretation [4]. AI has introduced groundbreaking improvements in the interpretation of X-rays, ultrasound images, magnetic resonance imaging (MRI), and computed tomography (CT) scans. Its impact extends beyond mere automation, fundamentally reshaping diagnostic processes to enhance speed, accuracy, and efficiency [5]. One of the main advantages of AI in imaging is its ability to significantly speed up analysis. Traditional methods, which are often slow and prone to human error, take a long time [6]. AI, on the other hand, can process and examine medical images in extremely short times, offering rapid diagnoses, a crucial aspect especially in emergency situations, where every second counts [7,8].

In addition to speed, AI also improves diagnostic accuracy. By analyzing vast medical datasets, algorithms can recognize patterns and abnormalities that might escape the human eye. This increased accuracy makes it possible to reduce diagnostic errors and ensure that patients receive appropriate and timely treatments [7].

In human medicine, the radiologists' error rate in interpreting diagnostic images has been thoroughly analyzed, considering the potential consequences for both patients and healthcare facilities [9,10]. Errors in the interpretation of diagnostic images can arise from various factors, such as missed lesions located in anatomically overlooked areas (commonly referred to as 'blind spots'), suboptimal use of imaging settings, overlapping pathological structures, or atypical presentations of certain diseases. These are common contributors that can compromise diagnostic accuracy [10]. Despite significant advancements in professional expertise, clinical knowledge, and technological innovation, the error rate in this field has remained surprisingly consistent over time, for example, staying below 15% in chest radiographic studies [10].

For this reason, several studies in human medicine suggest the use of AI-based softwares as support for the radiologist in image acquisition and interpretation [10,11]. However, in veterinary medicine, the error rate that characterizes the analysis of radiographic images has not been thoroughly studied [12,13]. Thus far, 18 studies have been published on the use of AI applied to the interpretation of chest radiographs (CXR) of dogs and cats. Of these, about six articles have focused on the detection of major alterations in cardiac silhouette. For example, in the study by Li and colleagues [14], they developed a convolutional neural network (CNN) based on the Visual Geometry Group 16 model to detect left atrium enlargement in lateral chest radiographs. The study used a database consisting of 792 radiographic images, which were then classified as "positive" or "negative" about left atrium enlargement. The images were subjected to classification by both a CNN and certified radiologists with the model achieved an overall accuracy of 82.71%, with a sensitivity of 68.42% and a specificity of 87.09%.

In research conducted by Fitzke [15], approximately 2.5 million thoracic and extrathoracic radiographs of dogs and cats were used to develop and train a DL model aimed at identifying various abnormalities. The results obtained proved to be quite encouraging, showing a low false positive rate (ranging from 0 to 0.057) and high sensitivity (ranging from 0 to 0.962). AI also presents significant limits that can be significant for its use in the daily routine. ML and DL models can be inaccurate due to issues with training data, such as insufficient or unrepresentative datasets, leading to false positives and negatives. Furthermore, the use of ML models on different patient populations from those used during the training phase may generate bias, considering different variables such as age, breed, morphology, differences in imaging techniques, and labeling modalities [16]. If such biased or inaccurate models are applied on a large scale, they could negatively affect the quality of care provided to many veterinary patients and mislead the vet.

A key issue is that AI algorithms are primarily trained and validated using large human medicine databases, whereas veterinary datasets are generally more limited [17].

Another ethical problem is that trust in artificial intelligence can decline, particularly with deep learning systems like CNNs, which, despite their accuracy, makes it difficult to understand what is behind AI decisions, such as the classification of radiological anomalies [18].

Beyond interpretability, ethical concerns also arise regarding patient data privacy in radiological AI models [17] and the environmental impact of energy-intensive AI systems [19].

Initially, we filtered the selected articles based on the anatomical domain, including only AI models that focused on CXR in both human and veterinary fields.

Then, this review adopted a critical and interpretative approach to analyze the most recent and cited applications of AI in the field of radiology, with a specific focus on comparing its use in the interpretation of CXR in human and veterinary medicine. Rather than strictly following a systematic review model, our aim was to offer a reasoned overview of currently available AI-based tools by selecting recent and clinically relevant studies in the context of diagnostics and patient care. The objective is to explore the similarities and differences between the two fields, both in terms of methodological approaches and performance, through a critical evaluation of the most recent literature. Particular attention was given to studies that not only present technical innovations but also offer practical insights into their applicability in real-world clinical settings. By identifying key trends, potential discrepancies in outcomes, and sector-specific challenges, this review aims to highlight the main differences among the latest publications in human and veterinary medicine. Ultimately, we seek to promote a collaborative and interdisciplinary understanding that can enhance the development of future studies in the field of AI, particularly within veterinary medicine.

2. Applications of Artificial Intelligence in Human Medicine: Chest Radiographs

Several recent studies have analyzed the performance of radiologists in interpreting diagnostic images, comparing the results obtained with and without AI assistance. For example, in the case of CXRs, the use of AI has been shown to significantly improve the

identification of abnormalities such as active tuberculosis [20], malignant nodules [21], and other pathologies of the greater thoracic district [22].

In this context, the AI system generates probabilities associated with specific diseases, sometimes accompanied by indications of the locations of abnormalities. Clinicians can supervise and integrate this information during or after image interpretation. A relevant example is the DL algorithm developed by Nam and colleagues [23], designed to detect 10 common abnormalities in CXR images, including pneumothorax, mediastinal thickening, pneumoperitoneum, nodules/masses, consolidation, pleural effusion, atelectasis, fibrosis, calcifications, and cardiomegaly. This algorithm also offers visual localization of abnormalities, providing additional support to physicians in the diagnostic process.

In the latter study [23], DLAD-10 (deep learning algorithm detecting) was trained with 146,717 radiographs of 108,053 patients using a ResNet34-based neural network with lesion-specific channels. As a yardstick, the performance of DLAD-10 was compared with a same-day CT-confirmed dataset (normal: abnormal) and an open-source dataset (PadChest; normal: abnormal) and accordingly compared with that of three radiologists. DLAD-10 correctly classified significantly more critical abnormalities (95.0% (57 out of 60)) than the pool of radiologists (84.4% (152 out of 180)). Ultimately, this DL algorithm showed excellent performance, improving radiologists' skills and reducing the time to report critical and urgent cases.

Similarly, Seah and colleagues in 2021 [24] conducted a study in which twenty radiologists reviewed CXR results of 127 clinical cases with and without the assistance of a DL algorithm and found that radiologists assisted by the DL algorithm showed much better reading performance, with higher areas under the curve (AUC) when assisted by the DL algorithm (AUC, 0.808; 95% confidence interval (CI)) than when not assisted (AUC, 0.713; 95% CI). The DL algorithm significantly improved radiologists' classification accuracy for 102 (80%) of the 127 clinical findings and was statistically non-inferior for 19 (15%) cases; moreover, none of these radiographic studies showed a decrease in accuracy when radiologists used the DL algorithms.

Furthermore, a more recent study by Banerjee et al. (2025) [25] reinforced the role of AI in radiology by demonstrating its effectiveness in cancer detection, particularly for chest radiographs. The study highlights AI's potential to assist radiologists in identifying lung nodules and other malignancies, which aligns with previous findings on the benefits of deep learning-assisted diagnoses [25].

However, while AI has demonstrated great promise, a recent study by Bottazzi, Ferrante, and Cheplygina (2024) [26] emphasized that the dataset used to train AI models significantly impacts their robustness and generalizability. Their study on the NIH CXR14 dataset highlights how different sources of data annotation affect model performance, raising important concerns regarding dataset bias in AI applications for chest radiographs.

Carloni and Tsaftaris (2025) [27] introduced a framework to enhance AI model robustness against domain shifts in chest radiographs. Their study demonstrated how domain variability significantly affects diagnostic reliability, reinforcing the need for domainadaptive AI.

Similarly, Liao and Xia (2025) [28] addressed label noise in CXR datasets by improving diagnostic accuracy despite noisy annotations. This research is particularly relevant to ensuring AI models remain reliable, even with imperfect training labels.

Additionally, Pedrosa et al. (2024) [29] presented an Anatomically-Guided Inpainting technique to reconstruct missing lung regions in CXR images. This approach allows AI models to handle incomplete scans effectively, making it particularly useful for real-world clinical applications.

The added value of AI assistance is particularly evident in specific situations such as emergencies. CXR is a simple and widely accessible imaging modality; however, its interpretation is not easy and often requires a high level of expertise and experience. Many studies have found substantial discrepancies in the interpretation of CXR images in the emergency department, ranging from 0.3% to 17% [30,31]. This type of discordant misinterpretation of critical cases could directly influence the clinical course and outcomes of patients. In addition, emergency room physicians often have little time or opportunity to reach an on-call radiologist for consultations [32]. Hwang [22] studied whether the application of a commercially available DL algorithm could improve physicians' reading performance for clinically relevant abnormalities on CXR scans in the context of emergency management. The assistance of the DL algorithm improved the sensitivity of radiology physicians' interpretation from 65.6% to 73.4%. Subsequently, in 2020, Kim and colleagues [33] reported that with DL algorithm support, physicians' diagnostic performance for pneumonia improved (sensitivity: 53.2% to 82.2%; specificity: 88.7% to 98.1%).

The study of Harmon [34] showed that a DL algorithm can achieve 90.8% accuracy with 84% sensitivity and 93% specificity in detecting COVID-19 pneumonia on CT scans. In addition, further studies have suggested that AI can support radiologists in distinguishing COVID-19 from other pulmonary infections on both CXR [20] and CT scans of the chest [34].

AI-based models have the potential to integrate multimodal data collected from patients, thus transforming the process of detection, diagnosis, and triage of suspected COVID-19 cases. In particular, the study by Ippolito [35] showed that AI can be extremely useful in daily clinical practice, especially in emergency departments where the large number of patients and the need for rapid responses can significantly affect diagnosis.

AI makes it possible to identify and distinguish different patterns of lung infections, improving clinical decision making, reducing response time, and increasing diagnostic confidence. As highlighted, AI systems can also detect characteristic signs of bacterial pneumonia, representing a crucial resource in settings where access to CT is limited or diagnostic expertise is insufficient.

The study also showed that AI can classify patients into three main categories: COVID-19 positive, pneumonia positive and healthy subjects, ensuring high accuracy and low error rates. These results help to strengthen the diagnostic confidence of radiologists, providing an effective tool to improve clinical care [35].

Particularly interesting is the 2024 study by Obuchowicz and colleagues [36] in which a novel radiographic motion simulation network was introduced that integrates U-Net with LSTM networks to simulate and predict respiratory motion of the lungs from singlephase chest radiographs. Then, a spatial transform network is applied for precise image deformation to reflect the real respiratory motion. The performance of the network is evaluated by qualitative and quantitative methods. This approach improves diagnostic capabilities by providing information on lung dynamics from static radiographs, offers a noninvasive alternative for lung function assessment, and increases diagnostic efficiency by extracting detailed information from routine chest radiographs.

A prospective, randomized, controlled study by Nam [37] involved 10,476 participants undergoing CXR during health checkups. Participants were randomly assigned to two groups: one using an AI-based system and one without AI. The primary objective was to measure the detection rate of treatable lung nodules calculated as the ratio of positively identified radiographs to total radiographs. Secondary objectives included rates of false reports, positive reports, and detection of malignant nodules and lung cancer. The detection rate of treatable lung nodules was significantly higher in the AI group (0.59% vs. 0.25%). The study highlights the potential of AI in improving the detection of pulmonary nodules in CXR scans during health screenings but emphasizes the need for further multicenter

research to confirm these findings and evaluate the clinical impact on a large scale [37]. However, a recent article was published by Garza-Frias [38] in which the use of an AI software-based (Qure.AI, Version 3.1.6) index to assess CXR images and predict the development of heart failure within one year of examination was investigated. A multicenter retrospective study of 1117 patients (mean age 67.6 years) with no pre-existing diagnoses of heart failure was conducted. Of these, 413 developed heart failure within one year of examination, while 704 did not. CXR images were analyzed with the qXR-HF (Qure.AI, Version 3.1.6) model, which provided information on cardiac silhouette, pleural effusion, and lung pattern. The results showed an AUC of 0.798 (95% CI 0.77–0.82), accuracy of 73%, sensitivity of 81%, and specificity of 68%. These data demonstrate the potential of opportunistic screening using AI in radiology, highlighting how automated analysis of CXR results can proactively identify patients at risk of developing heart failure, enabling timely and targeted interventions [38].

To address these issues, Park and Kooi (2024) [4] proposed Positive-Sum Fairness, an AI training methodology that ensures equitable diagnostic performance across demographic groups without degrading accuracy for any subgroup.

Furthermore, Queiroz et al. (2024) [39] introduced Backbone Foundation Models, which enable fairness evaluation in AI-based chest radiography without needing explicit demographic data. Their study suggests that AI fairness can be improved without compromising diagnostic accuracy.

3. Applications of Artificial Intelligence in Veterinary Medicine: Chest Radiographs

In human medicine, the use of AI is now a well-established reality, supported by a robust body of scientific literature. In contrast, in veterinary medicine, the exploration and publication of AI applications are still in their early stages. Nevertheless, some studies have begun investigating potential uses of AI in specialized fields, such as veterinary radiology.

In a study conducted by Yoon and colleagues [40], algorithms were employed to automatically analyze CXR images with the aim of detecting abnormalities in the cardiac silhouette, lung parenchyma, mediastinum, and pleural space, and distinguishing them from normal images. Initially, the study used models based on the Bag of Features (BOFs) technique, but the most promising results were achieved using CNNs. The CNNs reached accuracy levels ranging from 92.9% to 96.9% and sensitivities from 92.1% to 100%. In comparison, BOF models performed less effectively, with accuracy ranging between 79.6% and 96.9% and sensitivity between 74.1% and 94.8%. These findings highlight the superior performance of CNNs over BOF-based models.

Another interesting application of AI in veterinary radiology is found in the study by Kim and colleagues [41], who used commercial software (Vetology Innovations, San Diego, CA, USA) to detect cardiogenic pulmonary edema in 500 canine chest radiographs. Nineteen images were excluded due to technical issues or poor quality. Among the evaluated images, the system achieved an accuracy of 92.3%, with a specificity of 92.4% and a sensitivity of 91.3%. Notably, the negative predictive value (NPV) was 99%, indicating high reliability in confirming the absence of disease. However, the positive predictive value (PPV) was only 56%, suggesting that a positive result still requires a confirmation. It is also important to note that the number of radiographs used to train the system was significantly lower than the data volumes typically used in human medicine studies, and this could introduce bias.

The same software was later employed in two additional research studies. The first, conducted by Müller and colleagues [42], tested the software's ability to detect pleural effusion in radiographs from 62 dogs. The results showed an accuracy of 88.7%, sensitivity

of 90.2%, and specificity of 85.7%. However, it is worth mentioning that diagnoses were based solely on radiologist interpretation, without clinical, laboratory, or advanced imaging confirmation. This represents a limitation, as noted by the authors, considering that radiologists' accuracy in detecting pleural effusion can vary significantly, ranging from 67% to 92% [43]. As highlighted in the study, further research is needed to evaluate the true accuracy of this AI system [42].

An innovative approach was taken in the study by Pomerantz and colleagues [44] where the same AI software (Vetology Innovations, v. veterinary teleradiology) was used to detect pulmonary nodules and masses in 56 canine patients. The software's results were subsequently compared with computed tomography (CT) images of the same patients, which is considered the gold standard for detecting pulmonary nodules and masses [45]. The system achieved an accuracy of 69.3%, with a sensitivity of 55.4% and a specificity of 93.7%. While these performance metrics are lower than those seen in previous studies, they nonetheless confirm the potential of AI as a clinical support tool and diagnostic support. At the same time, they underscore the importance of combining AI with radiologist expertise. The decrease in performance compared to Müller's study [42] may be due to the use of CT as an objective diagnostic reference. Once again, the primary limitation was the small dataset used [44].

Currently, AI applications for interpreting feline chest radiographs remain limited, with only a few studies available, such as those by Banzato [46] and Dumortier [47]. Banzato's study involved training two deep neural network architectures, ResNet 50 and Inception V3, to recognize common thoracic radiographic findings in cats. These included bronchial patterns, pleural effusion, pulmonary masses, alveolar patterns, pneumothorax, cardiomegaly, and normal findings. The models showed good performance for most diagnostic categories, achieving area under the curve (AUC) values greater than 0.8. However, accuracy was lower for detecting cardiomegaly (AUC > 0.7) and particularly for pulmonary masses (AUC > 0.5).

In the second study, published by Dumortier and colleagues [47] in 2022, the ResNet50V2 neural network was used to classify feline thoracic images. A manual segmentation system was employed to define regions of interest. Although the results were promising, the method's effectiveness was limited by the small dataset (500 radiographs) and the requirement for human input during segmentation, both of which hinder its practical application in clinical settings.

As for the automatic analysis of the cardiac silhouette, veterinary literature is still quite sparse. One of the most significant studies was conducted by Burti and colleagues [48] who assessed the accuracy of four different CNN models in classifying the presence or absence of cardiomegaly using the Vertebral Heart Score (VHS) as a reference, taking breed variability into account. The best-performing model was based on the ResNet-101 architecture, achieving an AUC of 0.97. The study used a large dataset of 1465 lateral thoracic radiographs, the standard projection for VHS measurement, and concluded that this technology could serve as a valuable support tool for radiologists in clinical practice, especially given the somewhat subjective nature of this type of measurement.

In another study, researchers developed a DenseNet-121 neural network capable of automatically measuring VHS on lateral thoracic radiographs. Although the number of images used was limited (60 radiographs from dogs and cats), the system's measurements showed high concordance (>0.9) with those of two expert radiologists [49]. Again, a major limitation is that VHS measurements can vary depending on the operator. A similar investigation was conducted by Zhang and colleagues in 2021 [50], in which CNNs were trained to identify anatomical landmarks needed for VHS calculation. The system achieved an average accuracy of 91%, suggesting a promising clinical application to make VHS

assessment more objective and less susceptible to inter-operator variability, a common issue that can reduce the effectiveness of such measurements.

4. Discussion

The integration of AI into thoracic radiography, both in human and veterinary settings, represents a promising innovation that can help improve diagnostic quality and significantly reduce image interpretation time. Table 1 provides a summary of the main articles discussed in this review. However, it is critical to emphasize that AI systems should not be considered a substitute for the physician or radiologist. Although AI can provide valuable support, offering a "second opinion" and increasing diagnostic confidence, it remains susceptible to errors that only human clinical judgment can identify and contextualize. On the other hand, the study by Rudnay and Kovac (2024) [51] shows how the human factor, both psychological and physical, can lead to errors in interpretation, especially in the field of imaging.

Reference Task Species Banzato et al., 2021 [46] Detecting common radiographic findings Dog/cat Boissady et al., 2021 [49] Automatically measuring VHS Dog/cat Burti et al., 2020 [48] Classification of cardiomegaly based on VHS value Dog Detecting thoracic and extra-thoracic radiographic Fitzke et al., 2021 [15] Dog/cat abnormalities Garza-Frias et al., 2024 [38] Human Early detection of heart failure Identification of tuberculosis, malignant nodules, Hwang et al., 2019 [20] Human and other anomalies Hwang et al., 2019 [22] Human Use of commercial DL software in emergencies Human Ippolito et al., 2023 [35] Distinguish different patterns of lung infections Deep learning algorithm for detection of Kim et al., 2020 [33] Human pneumonia Presence/absence of cardiogenic pulmonary Kim et al., 2022 [41] Dog edema Li et al., 2020 [14] Detecting left atrial enlargement Dog Müller et al., 2022 [42] Presence of pleural effusion Dog Detection of 10 common abnormalities in CXR Nam et al., 2021 [23] Human scans Human Nam et al., 2023 [21] Detection of lung nodules Obuchowicz et al., 2024 [36] Real-time CXR Human Pomerantz et al., 2023 [44] Attendance of pulmonary nodules and masses Dog Effect of a comprehensive deep learning model on Seah et al., 2021 [24] Human the accuracy of CXR interpretation Normal vs. abnormal cardiac silhouette and Yoon et al., 2018 [40] Dog thoracic portions Zhang et al., 2021 [50] Identification of landmarks for calculating VHS Dog Human Banerjee et al., 2025 [25] AI in cancer diagnosis in radiology Human Juodelyte et al., 2024 [26] The importance of datasets

Table 1. Summary table of the main articles included in this review.

Since the human factor cannot be eliminated, neither in medicine nor in forensic practice, it is essential to recognize its limitations, and for this reason the integration of AI can provide valuable support [51].

The application of AI in medical imaging diagnostics could have negative effects on healthcare. For instance, a machine learning model designed to predict the risk of human pneumonia mistakenly assigned lower risk scores to asthma patients due to biases in the clinical data used for training. This error could potentially result in serious harm to patients [52].

The performance of AI systems hinges critically on their training foundation. Optimal results require large, well-structured datasets that precisely align with the intended diagnostic objectives—a significant challenge in veterinary medicine where such comprehensive data resources remain limited compared to human medical applications. Ensuring data consistency and representativeness can help mitigate biases and improve the reliability of AI-driven medical assessments.

In particular, the application of AI in veterinary medicine is still in its infancy and requires further study to fully assess its potential and adapt its techniques to the specific characteristics of species such as dogs and cats. An additional factor of complexity in veterinary medicine is the variability among different breeds, particularly in dogs, which can affect the accuracy of AI systems and lead to higher error rates. Therefore, it is essential that future research delves into these issues, encouraging increasingly effective and safe integration of these technologies into diagnostic practices, like what is being attempted in human medicine. One possible solution to overcome this limitation could be to begin studies focused on animal species with less phenotypic variability among breeds, such as cats. In this context, the relative homogeneity among feline breeds could facilitate the training of AI systems, and in this sense, it is comparable to the human case, reducing the complexity associated with analyzing more heterogeneous samples. Such an approach could represent a significant step toward the development of more accurate and generalizable AI models in veterinary medicine, opening new perspectives for the application of technology in the field. In this regard, in Figure 1 we depict potential AI tasks with a broad adoption in chest radiography analysis for human settings that could be useful to extend in the veterinary domain, such as heart segmentation and cardiac silhouette feature extraction through anatomical landmark detection.

Recent advances in landmark detection techniques for x-ray images have shown promising results that could be valuable for both human and veterinary applications. Di Via et al. [53] conducted a systematic study analyzing whether small-scale in-domain datasets provide any benefit for landmark detection over models pre-trained on large natural image datasets only. Their findings suggest that pre-training with ImageNet may be as effective as in-domain pre-training for anatomical landmark detection in x-ray images, which could simplify implementation in veterinary settings where large in-domain datasets are scarce. Furthermore, in a subsequent study, Di Via et al. [54] proposed a novel self-supervised pre-training approach using diffusion models for few-shot landmark detection in x-ray images, demonstrating good performance with as few as 50 annotated training images. These methodologies could be particularly valuable in veterinary medicine where annotated datasets are limited, especially for automated cardiac silhouette assessment as shown in Figure 1.

In this context, the guidance offered by studies conducted in human medicine is a valuable resource for advancing the veterinary field as well, providing models for research and application of artificial intelligence and other diagnostic technologies. However, as highlighted in this review, veterinary practice makes greater use of cardiac silhouette assessment in CXR images than in the human field, as advanced imaging techniques for

cardiac studies are less available. This approach, more widely adopted in the veterinary field, could prove to be an important opportunity for mutual growth and integration between the two fields. By fostering a bidirectional transfer of knowledge, more accurate and personalized diagnostic methods could be developed for both humans and animals. On the other hand, the use of AI is significantly more exploited in the interpretation of CXR scans during emergency situations in human medicine. This could serve as an excellent area to enhance tools in veterinary medicine, given the growing need to manage emergency cases effectively like in the study of Hwang and colleagues [22]. A crucial aspect that differentiates AI studies in human medicine from those in veterinary medicine is the size of the samples used. Studies in human medicine tend to benefit from significantly larger publicly available datasets, allowing for more robust statistical models and more accurate validation of results. In contrast, in veterinary medicine, the smaller sample size can affect the robustness of conclusions, posing a challenge for the effective application of AI. This underscores the importance of promoting cross-sectoral collaborations and developing strategies for augmenting and sharing datasets in the veterinary context as well, to bridge this disparity and maximize the potential of AI in both disciplines. Further studies are necessary to effectively implement the use of artificial intelligence in veterinary radiology, ensuring not only technical advancements but also clinically relevant applications.

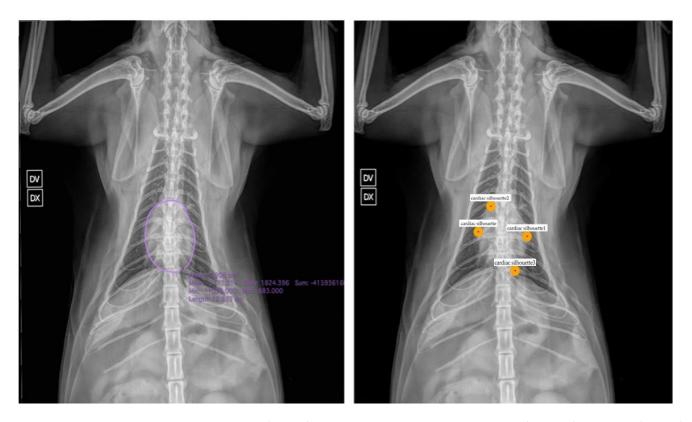


Figure 1. Potential AI tasks we envision on veterinary images, with examples on a cat chest radiograph. Left. Heart segmentation and extraction of metrics. Right. Anatomical landmark detection for analyzing cardiac silhouette. Output images and data from these tasks can potentially reveal information useful for automatic diagnosis in a classification framework.

5. Conclusions

The integration of artificial intelligence into thoracic radiography represents a promising innovation in both human and veterinary medicine, enhancing diagnostic quality and reducing interpretation time. However, AI cannot replace clinical judgment, as it is prone to biases and errors. In veterinary medicine, challenges arise from breed and species variability, as well as the limited availability of structured datasets. Studies on species with lower phenotypic variability, such as cats, could facilitate the training of more accurate models. In fact, in feline thoracic radiography, variations in thoracic conformation, cardiac silhouette appearance in different recumbency, tracheal size, and diaphragm shape are less pronounced among different breeds. Furthermore, promoting data sharing and integration is crucial for developing reliable AI applications in both fields. This review aimed to critically evaluate the current state and future potential of AI applications in thoracic radiography across human and veterinary domains, identifying both shared opportunities and discipline-specific challenges to guide future research and clinical implementation.

Author Contributions: Conceptualization, M.V. and A.R.; methodology, A.R. and R.D.V.; software, R.D.V., F.O. and V.P.P.; validation, M.V., F.O. and V.P.P.; resources, A.R and R.D.V.; data curation, A.R.; writing—original draft preparation, A.R. and R.D.V.; writing—review and editing, A.R., R.D.V., M.V., V.P.P. and F.O.; visualization, A.D.B., M.R., R.D.V., F.D.S., V.P.P., F.O., M.V. and A.R.; supervision, M.V., V.P.P. and F.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets presented in this article are not readily available due to restrictions from our Institutional Review Board.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Gottfredson, L.S. Mainstream science on intelligence: An editorial with 52 signatories, history and bibliography [Editorial]. Intelligence 1997, 24, 13–23. [CrossRef]
- Sung, J.J.; Stewart, C.L.; Freedman, B. Artificial intelligence in health care: Preparing for the fifth Industrial Revolution. *Med. J. Aust.* 2020, 213, 253–255.e1. [CrossRef] [PubMed]
- 3. Mitchell, T.M. Machine Learning; McGraw-Hill: New York, NY, USA, 1997.
- 4. Murphy, K.P. Machine Learning: A Probabilistic Perspective; MIT Press: Cambridge, MA, USA, 2012.
- Najjar, R. Redefining Radiology: A Review of Artificial Intelligence Integration in Medical Imaging. *Diagnostics* 2023, 13, 2760. [CrossRef]
- Hameed, B.M.Z.; Prerepa, G.; Patil, V.; Shekhar, P.; Zahid Raza, S.; Karimi, H.; Paul, R.; Naik, N.; Modi, S.; Vigneswaran, G.; et al. Engineering and clinical use of artificial intelligence (AI) with machine learning and data science advancements: Radiology leading the way for future. *Ther. Adv. Urol.* 2021, *13*, 17562872211044880. [CrossRef]
- Srivastav, S.; Chandrakar, R.; Gupta, S.; Babhulkar, V.; Agrawal, S.; Jaiswal, A.; Prasad, R.; Wanjari, M.B. ChatGPT in Radiology: The Advantages and Limitations of Artificial Intelligence for Medical Imaging Diagnosis. *Cureus* 2023, 15, e41435. [CrossRef]
- Pinto-Coelho, L. How Artificial Intelligence Is Shaping Medical Imaging Technology: A Survey of Innovations and Applications. Bioengineering 2023, 10, 1435. [CrossRef]
- 9. Siewert, B.; Sosna, J.; McNamara, A.; Raptopoulos, V.; Kruskal, J.B. Missed lesions at abdominal oncologic CT: Lessons learned from quality assurance. *Radiogr. A Rev. Publ. Radiol. Soc. N. Am. Inc.* **2008**, *28*, 623–638. [CrossRef]
- 10. Yun, S.J.; Kim, H.C.; Yang, D.M.; Kim, S.W.; Rhee, S.J.; Ahn, S.E. Diagnostic errors when interpreting abdominopelvic computed tomography: A pictorial review. *Br. J. Radiol.* **2017**, *90*, 20160928. [CrossRef]
- Degnan, A.J.; Ghobadi, E.H.; Hardy, P.; Krupinski, E.; Scali, E.P.; Stratchko, L.; Ulano, A.; Walker, E.; Wasnik, A.P.; Auffermann, W.F. Perceptual and Interpretive Error in Diagnostic Radiology-Causes and Potential Solutions. *Acad. Radiol.* 2019, 26, 833–845. [CrossRef]
- 12. Lamb, C.R.; Pfeiffer, D.U.; Mantis, P. Errors in radiographic interpretation made by veterinary students. *J. Vet. Med. Educ.* 2007, 34, 157–159. [CrossRef]

- 13. Cohen, J.; Fischetti, A.J.; Daverio, H. Veterinary radiologic error rate as determined by necropsy. *Vet. Radiol. Ultrasound Off. J. Am. Coll. Vet. Radiol. Int. Vet. Radiol. Assoc.* 2023, 64, 573–584. [CrossRef]
- Li, S.; Wang, Z.; Visser, L.C.; Wisner, E.R.; Cheng, H. Pilot study: Application of artificial intelligence for detecting left atrial enlargement on canine thoracic radiographs. *Vet. Radiol. Ultrasound Off. J. Am. Coll. Vet. Radiol. Int. Vet. Radiol. Assoc.* 2020, 61, 611–618. [CrossRef] [PubMed]
- Fitzke, M.; Stack, C.; Dourson, A.; Santana, R.M.B.; Wilson, D.; Ziemer, L.; Soin, A.; Lungren, M.P.; Fisher, P.; Parkinson, M. RapidRead: Global Deployment of State-of-the-Art Radiology AI for a Large Veterinary Teleradiology Practice. *arXiv* 2021, arXiv:2111.08165.
- 16. Coghlan, S.; Quinn, T. Ethics of using artificial intelligence (AI) in veterinary medicine. AI Soc. 2024, 39, 2337–2348. [CrossRef]
- Raymond Geis, J.; Brady, A.P.; Wu, C.C.; Spencer, J.; Ranschaert, E.; Jaremko, J.L.; Langer, S.G.; Kitts, A.B.; Birch, J.; Shields, W.F.; et al. Ethics of artificial intelligence in radiology: Summary of the joint European and North American multisociety statement. *Radiology* 2019, 293, 436–440. [CrossRef]
- 18. Quinn, T.P.; Jacobs, S.; Senadeera, M.; Le, V.; Coghlan, S. The three ghosts of medical AI: Can the black-box present deliver? *Artif. Intell. Med.* **2022**, 124, 102158. [CrossRef]
- 19. Crawford, K. *The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence;* Yale University Press: New Haven, CT, USA, 2021.
- Hwang, E.J.; Park, S.; Jin, K.N.; Kim, J.I.; Choi, S.Y.; Lee, J.H.; Goo, J.M.; Aum, J.; Yim, J.J.; Park, C.M.; et al. Development and Validation of a Deep Learning-based Automatic Detection Algorithm for Active Pulmonary Tuberculosis on Chest Radiographs. *Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am.* 2019, 69, 739–747. [CrossRef]
- Nam, J.G.; Park, S.; Hwang, E.J.; Lee, J.H.; Jin, K.N.; Lim, K.Y.; Vu, T.H.; Sohn, J.H.; Hwang, S.; Goo, J.M.; et al. Development and Validation of Deep Learning-based Automatic Detection Algorithm for Malignant Pulmonary Nodules on Chest Radiographs. *Radiology* 2019, 290, 218–228. [CrossRef]
- Hwang, E.J.; Park, S.; Jin, K.N.; Kim, J.I.; Choi, S.Y.; Lee, J.H.; Goo, J.M.; Aum, J.; Yim, J.J.; Cohen, J.G.; et al. Development and Validation of a Deep Learning-Based Automated Detection Algorithm for Major Thoracic Diseases on Chest Radiographs. *JAMA Netw. Open* 2019, 2, e191095. [CrossRef]
- 23. Nam, J.G.; Kim, M.; Park, J.; Hwang, E.J.; Lee, J.H.; Hong, J.H.; Goo, J.M.; Park, C.M. Development and validation of a deep learning algorithm detecting 10 common abnormalities on chest radiographs. *Eur. Respir. J.* 2021, *57*, 2003061. [CrossRef]
- 24. Seah, J.C.Y.; Tang, C.H.M.; Buchlak, Q.D.; Holt, X.G.; Wardman, J.B.; Aimoldin, A.; Esmaili, N.; Ahmad, H.; Pham, H.; Lambert, J.F.; et al. Effect of a comprehensive deep-learning model on the accuracy of chest x-ray interpretation by radiologists: A retrospective, multireader multicase study. *Lancet Digit. Health* **2021**, *3*, e496–e506. [CrossRef] [PubMed]
- 25. Banerjee, A.; Shan, H.; Feng, R. Editorial: Artificial intelligence applications for cancer diagnosis in radiology. *Front. Radiol.* **2025**, *5*, 1493783. [CrossRef] [PubMed]
- Juodelyte, D.; Lu, Y.; Jiménez-Sánchez, A.; Bottazzi, S.; Ferrante, E.; Cheplygina, V. Source Matters: Source Dataset Impact on Model Robustness in Medical Imaging. In *Applications of Medical Artificial Intelligence. AMAI 2024*; Wu, S., Shabestari, B., Xing, L., Eds.; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2025; Volume 15384.
- Carloni, G.; Tsaftaris, S.A.; Colantonio, S. CROCODILE: Causality Aids RObustness via COntrastive DIsentangled LEarning. In Uncertainty for Safe Utilization of Machine Learning in Medical Imaging. UNSURE 2024; Sudre, C.H., Mehta, R., Ouyang, C., Qin, C., Rakic, M., Wells, W.M., Eds.; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2025; Volume 15167.
- Ai, X.; Liao, Z.; Xia, Y. GLANCE: Combating Label Noise Using Global and Local Noise Correction for Multi-label Chest X-Ray Classification. In *Uncertainty for Safe Utilization of Machine Learning in Medical Imaging*. UNSURE 2024; Sudre, C.H., Mehta, R., Ouyang, C., Qin, C., Rakic, M., Wells, W.M., Eds.; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2025; Volume 15167.
- Pedrosa, J.; Pereira, S.C.; Silva, J.; Mendonça, A.M.; Campilho, A. Anatomically-Guided Inpainting for Local Synthesis of Normal Chest Radiographs. In *Deep Generative Models. DGM4MICCAI* 2024; Mukhopadhyay, A., Oksuz, I., Engelhardt, S., Mehrof, D., Yuan, Y., Eds.; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2025; Volume 15224.
- Brunswick, J.E.; Ilkhanipour, K.; Seaberg, D.C.; McGill, L. Radiographic interpretation in the emergency department. *Am. J. Emerg. Med.* 1996, 14, 346–348. [CrossRef] [PubMed]
- 31. Preston, C.A.; Marr, J.J., III; Amaraneni, K.K.; Suthar, B.S. Reduction of "callbacks" to the ED due to discrepancies in plain radiograph interpretation. *Am. J. Emerg. Med.* **1998**, *16*, 160–162. [CrossRef]
- 32. Gatt, M.E.; Spectre, G.; Paltiel, O.; Hiller, N.; Stalnikowicz, R. Chest radiographs in the emergency department: Is the radiologist really necessary? *Postgrad. Med. J.* 2003, *79*, 214–217. [CrossRef]

- Kim, J.H.; Kim, J.Y.; Kim, G.H.; Kang, D.; Kim, I.J.; Seo, J.; Andrews, J.R.; Park, C.M. Clinical Validation of a Deep Learning Algorithm for Detection of Pneumonia on Chest Radiographs in Emergency Department Patients with Acute Febrile Respiratory Illness. J. Clin. Med. 2020, 9, 1981. [CrossRef]
- Harmon, S.A.; Sanford, T.H.; Xu, S.; Turkbey, E.B.; Roth, H.; Xu, Z.; Yang, D.; Myronenko, A.; Anderson, V.; Amalou, A.; et al. Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets. *Nat. Commun.* 2020, 11, 4080. [CrossRef]
- 35. Ippolito, D.; Maino, C.; Gandola, D.; Franco, P.N.; Miron, R.; Barbu, V.; Bologna, M.; Corso, R.; Breaban, M.E. Artificial Intelligence Applied to Chest X-ray: A Reliable Tool to Assess the Differential Diagnosis of Lung Pneumonia in the Emergency Department. *Diseases* **2023**, *11*, 171. [CrossRef]
- 36. Obuchowicz, R.; Strzelecki, M.; Piórkowski, A. Clinical Applications of Artificial Intelligence in Medical Imaging and Image Processing-A Review. *Cancers* 2024, *16*, 1870. [CrossRef]
- Nam, J.G.; Hwang, E.J.; Kim, J.; Park, N.; Lee, E.H.; Kim, H.J.; Nam, M.; Lee, J.H.; Park, C.M.; Goo, J.M. AI Improves Nodule Detection on Chest Radiographs in a Health Screening Population: A Randomized Controlled Trial. *Radiology* 2023, 307, e221894. [CrossRef]
- Garza-Frias, E.; Kaviani, P.; Karout, L.; Fahimi, R.; Hosseini, S.; Putha, P.; Tadepalli, M.; Kiran, S.; Arora, C.; Robert, D.; et al. Early Detection of Heart Failure with Autonomous AI-Based Model Using Chest Radiographs: A Multicenter Study. *Diagnostics* 2024, 14, 1635. [CrossRef] [PubMed]
- Queiroz, D.; Anjos, A.; Berton, L. Using Backbone Foundation Model for Evaluating Fairness in Chest Radiography Without Demographic Data (FAIMI 2024, EPIMI 2024). In *Ethics and Fairness in Medical Imaging.*; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2025; Volume 15198.
- 40. Yoon, Y.; Hwang, T.; Lee, H. Prediction of radiographic abnormalities by the use of bag-of-features and convolutional neural networks. *Vet. J.* **2018**, 237, 43–48. [CrossRef] [PubMed]
- Kim, E.; Fischetti, A.J.; Sreetharan, P.; Weltman, J.G.; Fox, P.R. Comparison of artificial intelligence to the veterinary radiologist's diagnosis of canine cardiogenic pulmonary edema. *Vet. Radiol. Ultrasound Off. J. Am. Coll. Vet. Radiol. Int. Vet. Radiol. Assoc.* 2022, 63, 292–297. [CrossRef] [PubMed]
- Müller, T.R.; Solano, M.; Tsunemi, M.H. Accuracy of artificial intelligence software for the detection of confirmed pleural effusion in thoracic radiographs in dogs. *Vet. Radiol. Ultrasound Off. J. Am. Coll. Vet. Radiol. Int. Vet. Radiol. Assoc.* 2022, 63, 573–579. [CrossRef]
- Cicero, M.; Bilbily, A.; Colak, E.; Dowdell, T.; Gray, B.; Perampaladas, K.; Barfett, J. Training and Validating a Deep Convolutional Neural Network for Computer-Aided Detection and Classification of Abnormalities on Frontal Chest Radiographs. *Investig. Radiol.* 2017, 52, 281–287. [CrossRef]
- Pomerantz, L.K.; Solano, M.; Kalosa-Kenyon, E. Performance of a commercially available artificial intelligence software for the detection of confirmed pulmonary nodules and masses in canine thoracic radiography. *Vet. Radiol. Ultrasound Off. J. Am. Coll. Vet. Radiol. Int. Vet. Radiol. Assoc.* 2023, 64, 881–889. [CrossRef]
- Diederich, S.; Semik, M.; Lentschig, M.G.; Winter, F.; Scheld, H.H.; Roos, N.; Bongartz, G. Helical CT of pulmonary nodules in patients with extrathoracic malignancy: CT-surgical correlation. *Am. J. Roentgenol.* 1998, 172, 353. [CrossRef]
- 46. Banzato, T.; Wodzinski, M.; Tauceri, F.; Donà, C.; Scavazza, F.; Müller, H.; Zotti, A. An AI-Based Algorithm for the Automatic Classification of Thoracic Radiographs in Cats. *Front. Vet. Sci.* **2021**, *8*, 731936. [CrossRef]
- 47. Dumortier, L.; Guépin, F.; Delignette-Muller, M.L.; Boulocher, C.; Grenier, T. Deep learning in veterinary medicine, an approach based on CNN to detect pulmonary abnormalities from lateral thoracic radiographs in cats. *Sci. Rep.* **2022**, *12*, 11418. [CrossRef]
- 48. Burti, S.; Longhin Osti, V.; Zotti, A.; Banzato, T. Use of deep learning to detect cardiomegaly on thoracic radiographs in dogs. *Vet. J.* **2020**, *262*, 105505. [CrossRef]
- Boissady, E.; De La Comble, A.; Zhu, X.; Abbott, J.; Adrien-Maxence, H. Comparison of a Deep Learning Algorithm vs. Humans for Vertebral Heart Scale Measurements in Cats and Dogs Shows a High Degree of Agreement Among Readers. *Front. Vet. Sci.* 2021, *8*, 764570. [CrossRef] [PubMed]
- 50. Zhang, M.; Zhang, K.; Yu, D.; Xie, Q.; Liu, B.; Chen, D.; Xv, D.; Li, Z.; Liu, C. Computerized assisted evaluation system for canine cardiomegaly via key points detection with deep learning. *Prev. Vet. Med.* **2021**, *193*, 105399. [CrossRef] [PubMed]
- 51. Rudnay, M.; Kováč, P. Bias, fatigue and other factors as potential source of errors in medical practice and forensic medicine. *Rom. Soc. Leg. Med.* **2024**, *32*, 46–51.
- 52. Caruana, R.; Lou, Y.; Gehrke, J.; Koch, P.; Sturm, M.; Elhadad, N. Intelligible Models for HealthCare: Predicting Pneumonia Risk and Hospital 30-day Readmission. In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '15), Sydney, Australia, 10–13 August 2015; Association for Computing Machinery: New York, NY, USA, 2015; pp. 1721–1730. [CrossRef]

- Di Via, R.; Santacesaria, M.; Odone, F.; Pastore, V.P. Is In-Domain Data Beneficial in Transfer Learning for Landmarks Detection in X-Ray Images? In Proceedings of the IEEE International Symposium on Biomedical Imaging, ISBI 2024, Athens, Greece, 27–30 May 2024; pp. 1–5. [CrossRef]
- Di Via, R.; Odone, F.; Pastore, V.P. Self-Supervised Pre-Training with Diffusion Model for Few-Shot Landmark Detection in X-Ray Images. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, WACV 2025, Tucson, AZ, USA, 26 February–6 March 2025; pp. 3886–3896. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.