Final Report 5ARIP10

Lemar Abdi
Department of Electrical Engineering
1463225
TU Eindhoven
l.abdi@student.tue.nl

Halil Akdemir
Department of Electrical Engineering
1280376
TU Eindhoven
h.i.akdemir@student.tue.nl

Thodoris Chronopoulos
Department of Electrical Engineering
1886312
TU Eindhoven
t.chronopoulos@student.tue.nl

***Abstract*—This report discusses the development of a deep learning algorithm for real-time denoising of low-dose X-ray images in C-arms, which will be translated into a start-up. The goal is to reduce radiation exposure for patients and especially medical professionals while maintaining image quality and diagnostic information. A methodology is proposed that involves a literature review, market research, a proof-of-concept phase using maturity steps, testing on real scans, and verification and validation of the technology. Data acquisition, curation, storage and management as well as the business case for thee proposed start-up are discussed. A back-of-the-envelope calculation is performed and shows significant market size and revenue for the start-up. We have chosen to use a KBNet (Kernel Basis Network for Image Restoration) for the denoising task. Experimental results using a deep learning model with pre-trained weights have shown promising outcomes, with untapped potential for further improvement through transfer learning techniques. The model evaluated on the test dataset (VinDr- SpineXR dataset) gives a PSNR (Peak signal-to-noise ratio) of 82.25 dB and a MSE score of 0.0004. Furthermore, stacking grey-scale images along the input channel dimension utilized temporal information and improves the model performance on video.**

I. Introduction

C-arm machines are devices that use X-rays to produce real-time images of the anatomy and allow for precise guidance of instruments and devices. In this essay, we will discuss x-ray image denoising for C-arm radiation reduction. C-arm systems are widely used for fluoroscopic guidance during interventional procedures such as angiography, endoscopy, biopsy, and orthopedic surgery. However, C-arm systems also expose both patients and physicians to relatively low doses of radiation. Unlike patients, who receive a single dose of radiation during the procedure, physicians are repeatedly exposed to radiation over time, which acculumates to a high radiation exposure; increasing their cumulative risk. X-rays are associated with potential health risks such as cancer and tissue damage, especially for repeated or prolonged exposures [1]. Consequently, the maximum permissible dose of radiation exposure for occupational workers in the USA is 50 millisieverts (mSv) per year [2]. While the European Union (EU) has set a limit of 20 mSv per year for occupational exposure [3]. In the paper "Radiation Exposure of the Hand and Chest during C-arm Fluoroscopy-Guided Procedures" by Jung et al. [4] it has been found that the average radiation exposure to the physician's hand was 0.13 mSv per procedure. Therefore, it is important to reduce the x-ray dose as much as possible while maintaining the image quality and diagnostic information.

One way to achieve this goal is to apply denoising techniques to x-ray images that are acquired with low dose levels. Denoising techniques aim to remove or reduce the noise that corrupts the image signal, which is mainly due to the quantum nature of x-ray photons and the electronic noise of the detector. However, denoising x-ray images is challenging because of the mixed signal-dependent noise characteristics, the high dynamic range of the image intensities, and the presence of fine structures and edges that are essential for diagnosis.

In recent years, deep learning based denoising techniques have emerged as a promising approach to address these challenges. CNN-based architectures are trained on large datasets of noisy and clean images or image patches. These models can learn complex and nonlinear mappings between noisy and clean images and can adapt to different noise levels and image contents. Learning-based techniques can be divided into two categories: supervised and unsupervised. Supervised techniques require pairs of noisy and clean images for training, while unsupervised techniques only need noisy images. Research has been done on deep learning techniques for low-dose CT or X-ray image denoising [5]–[8]. However, we as a start-up will propose the novel idea of implementing deep learning techniques for real-time denoising of C-arm X-ray images that will result in a lowered radiation dosage.

Philips Healthcare, Siemens Healthineers, Canon Medical Systems, and GE Healthcare, are all leading players in the medical imaging industry. Philips Healthcare's ClarityIQ [9] reduces noise and artifacts in X-ray images that can be used for X-ray dose management for the patient. Siemens Healthineers' AI Rad Companion [10], Canon Medical Systems' CXDI-RF for the Celex [11] and GE Healthcare's Critical Care Suite [12] all aim to provide high image quality for X-rays. However, these products do not utilize deep learning methods, do not aim to lower radiation dosages or do not focus on the C-arm. Thus, the proposed proprietary innovation is unique and is not offered as a product by the key players in the medical imaging industry.

II. Methodology

Our project aims to develop a deep learning algorithm using deep learning for low-dose X-ray images in C-arms. This algorithm will enable the denoising of low-dose images in real-time, resulting in reduced radiation exposure for both patients and medical professionals. To achieve this, we will follow a methodology that involves several key steps.

First, a literature review has been done to gather information on the latest developments in deep learning, denoising, and low-dose X-ray imaging. Furthermore, market research has been done to identify competitors, and opportunities and to perform a back-of-the-envelope calculation. This will provide a solid foundation for our project and help us to identify the most promising approaches to achieving our goal.

Next, we will develop a proof-of-concept algorithm using maturity steps in data, which will enable us to test and refine our approach before moving on to real-time X-ray scans. This phase will involve the development and training of deep learning models, as well as the optimization of the architecture and algorithm.

Once we have validated our proof-of-concept algorithm, we will move on to testing our approach on real-time X-ray scans, eventually on a real C-arm. We will work with medical professionals to ensure that our algorithm meets the necessary standards for image quality and radiation exposure reduction.

Throughout the project, we will employ an agile methodology that emphasizes frequent communication and collaboration. We will hold regular meetings to review progress, discuss challenges, and adjust our approach as needed. We will also use project management software to keep track of tasks, deadlines, and progress, ensuring that we stay on track and meet our goals.

Finally, we will evaluate the effectiveness of our algorithm using both quantitative and qualitative measures. The software will be verified to determine the quality and also validated to ensure that the requirements are met. The resulting images are assessed in the quality of the reconstruction. We will also seek feedback from medical professionals to ensure that our approach is practical and useful in a real-world setting.

In summary, our methodology involves a thorough review of the literature, a proof-of-concept phase using maturity steps, testing on real scans, and evaluation of the effectiveness of our approach. Through collaboration, communication, and agile project management, we are confident that we will be able to develop a highly effective real-time denoising algorithm using deep learning for low-dose X-ray images.

III. Algorithm Development

Our projects aims to develop an algorithm that can reconstruct images taken by a C-arm, while lowering the dosage and maintaining the same or better image quality. In order to do this we first have to look at the effect of lowering the dosage. When this is known we will look into methods to undo these effects in order to maintain our image quality while lowering the dosage.

A. Noise caused by limited radiation

Lowering the radiation dose in medical imaging can result in images with different types of noise. The two main types of noise in medical imaging are white noise and quantum noise. When the radiation dosage is lowered to reduce the patient's exposure to radiation, the number of photons detected by the X-ray detector is also reduced. As a result, the photon statistics become a more dominant source of noise, which can lead to increased image noise.

When initially lowering the radiation, white noise will appear [13]. Only when lowering the dosage under an extreme threshold, quantum noise will start top appear [14], that is why we only focus on the white noise.

An example of a benchmark image with different types of noise can be seen in figure 1.

B. Limited radiation through sparse sampling

Sparse sampling in 2D X-ray imaging can be achieved by placing a grid in front of the X-ray emitter to limit the number of X-rays that pass through the object being imaged. This results in lower radiation dose to the patient and reduced image noise. The grid consists of thin lead strips or other radiation-absorbing materials that are arranged in a parallel configuration. The X-rays that pass through the gaps between the strips are used to form the image, while those that are absorbed by the strips are discarded. This will lead to some kind of noise pattern in the images. The pattern will depend on the kind of grid is used.

Sparse sampling has shown to be a good solution in other types of medical imaging like CT [15]. In CT, the reason why sparse sampling works is due to the fact that the rotations and occlusions of the object being scanned occur at different angles, allowing for different areas of the object to be captured in different projections. This enables the reconstruction algorithm to fill in the missing data and produce a high-quality image.

However, when a grid is placed in front of a 2D x-ray, the occlusions will not change position and will be present in every projection. This means that there will be less variation in the captured data, and the reconstruction algorithm will have a harder time filling in the missing information. As a result, the quality of the denoised image will not be as good as in CT.

Another downside of using a grid is introducing a mechanical aspect. The grid will have to be designed and mounted to the x-ray emitter, the will also introduce compatibility issues since the x-ray emitter is not universal. In order for our startup to keep the problem as one dimensional as possible, this sparse sampling approach was disregarded.

C. Deep learning based methods

There are various techniques that can be used to mitigate the effects of noise in medical images. These include hardware-based methods such as optimizing the acquisition parameters, using special detectors or filters, and post-processing methods such as denoising and deconvolution techniques.

Deep learning-based methods have gained significant attention in recent years for their ability to learn complex mappings between the noisy and noise-free images. These methods

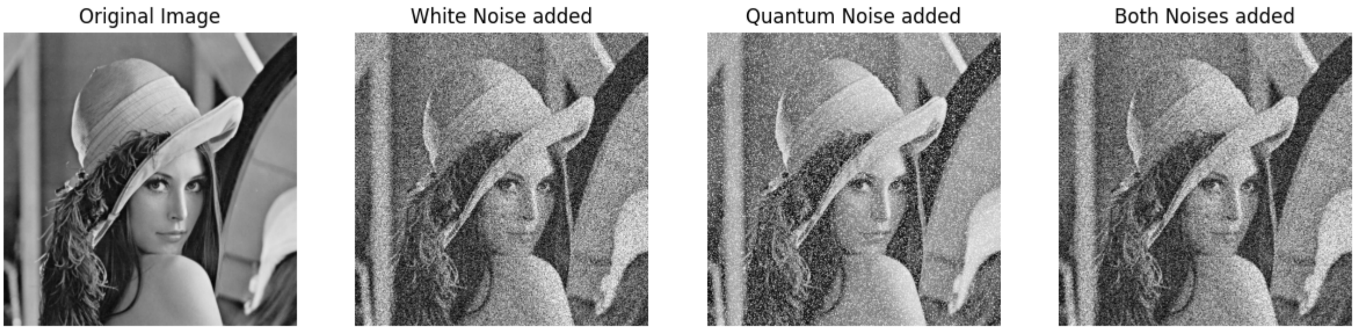

Fig. 1: Different noise types applied onto an original and lower contrast image

have shown promising results in various medical imaging applications, including denoising CT, MRI, and X-ray images. The use of deep learning-based methods has shown significant improvements over traditional denoising techniques in terms of denoising performance and preserving image details.

One of the popular deep learning-based denoising techniques is the use of convolutional neural networks (CNNs). These networks can be trained using a large dataset of noisy and noise-free images to learn a mapping between the two. The trained network can then be used to denoise new images. CNNs can capture spatial information in images, making them well-suited for tasks like image classification and segmentation. But, they may not always perform well for tasks that require modeling of global context, such as image denoising.

Autoencoders are a type of neural network that can be used for unsupervised learning. They are commonly used for image denoising and restoration tasks. Autoencoders consist of an encoder network that maps the input image into a low-dimensional latent space, and a decoder network that maps the latent representation back to the original image space. The encoder and decoder networks are trained together to minimize the reconstruction error between the input and output images. Autoencoders can be effective at removing noise from images while preserving important features, making them useful for tasks like medical image analysis. However, autoencoders can be sensitive to overfitting, especially when the training dataset is small or the model is too complex.

One popular autoencoder architecture for medical image denoising is the U-Net [16]. The U-Net architecture was originally proposed for image segmentation tasks, but it has since been adapted for various other image-related tasks, including image denoising. The U-Net architecture consists of a contracting path that captures the context of the input image and a symmetric expanding path that produces the denoised output.

Another type of architecture that can be used are transformers. Transformers are a type of neural network architecture that have recently gained popularity in the field of natural language processing. However, they have also shown promising results in various computer vision tasks, including image denoising [17].

One advantage of transformers is their ability to capture long-range dependencies in the input data. This makes them particularly effective for tasks that require modeling of global context, such as image denoising. Transformers also have a parallelizable architecture, which makes them efficient for processing large datasets.

However, transformers can be computationally expensive and require large amounts of memory. This can make them challenging to train on limited computational resources. Additionally, transformers may not perform as well as convolutional neural networks (CNNs) for tasks that require spatial information, such as image segmentation.

D. Performance metrics

Assessing the performance of denoising models is a crucial step in evaluating their effectiveness for a given task. There are several performance metrics that can be used to evaluate denoising models, depending on the specific requirements of the task.

One commonly used metric is peak signal-to-noise ratio (PSNR), which measures the ratio between the maximum possible value of a signal and the amount of noise present. Higher PSNR values indicate better denoising performance. However, PSNR does not always correlate well with perceived image quality, and may not be the best metric for tasks such as medical image analysis where visual clarity is critical.

Another widely used metric is mean squared error (MSE), which measures the average squared difference between the denoised image and the ground truth image. While MSE is a simple and intuitive metric, it can be sensitive to outliers and may not accurately capture differences in perceived image quality.

When selecting a denoising model based on performance metrics, it is important to consider both the accuracy and efficiency of the model. Deep learning models that achieve high denoising performance may be computationally expensive and require significant amounts of training data. Thus, it is important to consider trade-offs between model accuracy and computational efficiency when selecting a denoising model.

In conclusion, the selection of denoising models should be based on careful consideration of the task requirements, available training data, and performance metrics. While metrics such as PSNR and MSE are commonly used, it is important to evaluate the performance of models using multiple metrics, including those that more accurately capture perceived image quality.

E. KBNet

Based on the metrics above a KBNet (Kernel Basis Network for Image Restoration) [18] was chosen to do the denoising task. A huge advantage of this model is that it implements state-of-the-art (SOTA) methods like attention mechanisms to perform as good as possible while keeping the computational cost relatively low.

KBNet is a novel neural network architecture for image restoration tasks, such as denoising, deraining, and deblurring. It is based on the idea of kernel basis attention (KBA), which allows the network to learn adaptive spatial aggregation weights from learnable kernel bases. KBNet outperforms previous state-of-the-art methods on several benchmarks while being more efficient in terms of computation and memory.

The model is pretrained on the Smartphone Image Denoising Dataset (SSID) [19], which is a standard benchmark dataset for image denoising. The model is trained on images containing Gaussian noise which is relevant because our goal is also to denoise Gaussian noise. An example of image denoising using KBNet is shown in figure 2.

F. Simulation

To investigate the denoising of medical images, particularly focusing on whole image sequences, we conducted a simulation to progressively address the denoising problem. Our simulation consisted of two stages: denoising a single image and denoising a sequence of images. In the initial stage, we focused on denoising a single grayscale medical image. This served as a foundation for understanding the model's performance in removing noise from individual frames. By applying the image denoising model to each frame independently, we assessed its capability to reduce noise and preserve critical details within a single image. After single image denoising, we will progress to the denoising of image sequences. At this stage we also added a moving black line to simulate a needle during surgery, which can be seen in figure 3. This stage involved applying the denoising model to each frame in a sequential manner. However since the frames are all denoised separately the spatial information in the sequence is not considered. To exploit spatial information and potentially improve denoising performance further, we investigated the stacking of grayscale images into multi-layer representations. By stacking the grayscale images into 3-layer images, we aimed to incorporate spatial correlations across adjacent frames. This approach allowed the model to consider both temporal and spatial information during the denoising process, potentially enhancing noise reduction and image quality.

G. Deployment

When it comes to deploying deep learning models, there are two main options: cloud computing or edge computing.

The choice between edge computing and cloud computing depends on the trade-off between network bandwidth, latency, and complexity. Different applications and scenarios may have different preferences and constraints on these factors. For example, a study by IBM [20] proposed a framework to evaluate the cost-effectiveness of edge computing and cloud computing based on four criteria: data volume, data velocity, data variety, and data value. The study suggested that edge computing is more suitable for applications that have high data volume, high data velocity, low data variety, and low data value, whereas cloud computing is more suitable for applications that have low data volume, low data velocity, high data variety, and high data value.

For the task of real-time X-ray image denoising in clinical applications, we can analyze which categories of the four criteria it falls into:

(i) Data Volume: X-ray images typically have a moderate to high data volume. They contain a significant amount of image and meta data, which can contribute to a larger data volume.
(ii) Data Velocity: The data velocity for real-time X-ray image denoising is generally high. In clinical applications, the X-ray images need to be processed quickly and efficiently to make it real-time.
(iii) Data Variety: X-ray images have a relatively low data variety. Unlike other types of medical imaging data that may have multiple dimensions, channels, or modalities, X-ray images are typically grayscale two-dimensional images, which simplifies the data variety aspect.
(iv) Data Value: X-ray images hold high data value in clinical applications. They provide crucial information for diagnosing various medical conditions and guiding treatment decisions. The accuracy and quality of X-ray images play a vital role in ensuring effective clinical outcomes.

Based on these considerations, real-time X-ray image denoising in clinical applications falls into the following categories: high data velocity, low data variety, moderate data volume and high data value. Therefore, edge computing will be used.

IV. DATA ACQUISITION

A. Maturity steps

The primary goal of this project is to develop an algorithm capable of denoising low-dose radiation x-ray images in real-time, with the ultimate aim of improving the accuracy and efficiency of medical diagnoses while reducing the harmful radiation for the patients and practitioners. The four maturity steps have been defined to test and validate the methodology, which includes evaluating the algorithm's performance in handling different types of noise and medical imaging-specific noise. The project's ultimate goal is to improve medical imaging procedures and advance medical diagnosis and treatment.

The first step involves evaluating the denoising model with the CIFAR-10 dataset [21] to establish a baseline performance. This will help determine the effectiveness of the algorithm in comparison to existing denoising techniques.

Fig. 2: Image denoising through KBNet

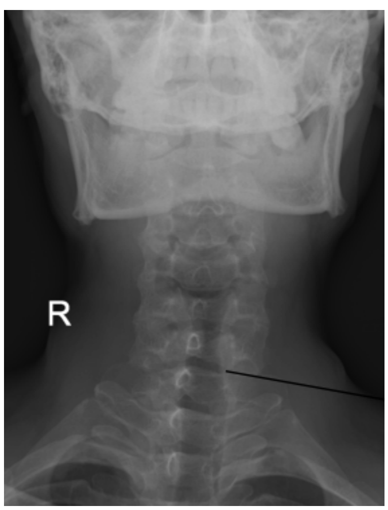

Fig. 3: A frame out of the simulated video sequence: the black line simulating a needle can be seen

The second step involves using the VinDr-SpineXR dataset [22] to perform 2D denoising on x-ray images of the spine by adding artificial noise. This step will enable the methodology to be validated and the algorithm's effectiveness in denoising medical images to be evaluated.

The third step involves applying white noise to the CIFAR-10 dataset to develop a prototype that can handle this type of noise. However, we will focus on pushing the white noise to extreme levels to simulate low X-ray radiation levels. This will test the algorithm's ability to denoise X-ray images in scenarios where the radiation levels are significantly reduced, leading to noisy and low-quality images.

In the fourth phase, we will continue to use the VinDr-SpineXR dataset to perform 2D denoising on X-ray images of the spine. However, just like in the third maturity step, we will simulate low X-ray radiation levels by adding extreme levels of white noise. This step will enable the evaluation of the algorithm's effectiveness in handling medical imaging-specific noise that is caused by reduced radiation levels, mimicking scenarios where the images are of poor quality due to limited exposure.

Completing these four maturity steps will enable the algorithm's performance to be evaluated and its suitability for denoising x-ray images to be determined. The project has the potential to improve medical imaging procedures and advance medical diagnosis and treatment.

B. Data

In the first and third maturity steps, the CIFAR-10 dataset [23] will be used. The CIFAR-10 dataset is a well-known benchmark dataset in computer vision, consisting of 60,000 32x32 color images in 10 different classes. However, since medical X-rays are grayscale images, we will convert the CIFAR-10 dataset to grayscale as well. This adaptation allows us to better simulate the characteristics of X-ray images and ensure that the denoising algorithm performs optimally in a grayscale context. Furthermore, working with grayscale images also offers the advantage of reduced computational complexity and resource requirements compared to color images.

The second and fourth maturity steps of this research project involve utilizing the VinDr-SpineXR dataset [24], which is a comprehensive medical image dataset specifically designed for detecting and classifying spinal lesions from radiographs. This dataset comprises over 10,000 X-ray images of the spine, along with annotations that indicate the location and type of spinal lesion present. The VinDr-SpineXR dataset was created by VinBrain, a company specializing in developing artificial intelligence solutions for healthcare, in collaboration with radiologists and medical professionals. Various studies have already leveraged this dataset to develop deep-learning algorithms for spinal lesion detection [25] [24] [26]. The high quality of this dataset makes it suitable for our case. Furthermore, by training and evaluating the denoising algorithm on the VinDr-SpineXR dataset, which has established credibility and reliability, we can expect the algorithm to deliver a superior level of quality and performance.

C. Data de-identification

The VinDr-SpineXR dataset has undergone de-identification, with the removal of protected health information (PHI) and retention of only basic demographic information, namely sex, and age, or parameters for image processing through the use of DICOM tags. However, since these DICOM tags are not relevant to the denoising algorithm that is being developed, they will also be removed. As a result, the dataset will be completely anonymous, further safeguarding the privacy of the individuals whose data is included. This approach is in line with GDPR, which requires the protection of personal data and will ensure that the dataset can be used for research purposes in a manner that is both ethical and legal.

D. Data storage and management

To guarantee the security and integrity of both the VinDr-SpineXR dataset and the simulated data created through the C-arm simulation, it is crucial to implement a combination of measures to prevent unauthorized access and data loss. Access controls will be established to limit dataset access to authorized individuals only. This includes using strong passwords or passphrases, multi-factor authentication, and access restrictions based on roles and responsibilities to ensure that only those with the necessary permissions can access the datasets.

Furthermore, regular vulnerability management practices will be implemented to identify and mitigate potential security threats to both datasets. This includes performing regular vulnerability scans and promptly patching any systems with known vulnerabilities. Keeping up to date with the latest security patches and updates will help reduce the risk of data breaches and other security incidents.

In addition to access controls and vulnerability management, both datasets will be regularly backed up to ensure that they can be restored in the event of data loss due to accidental deletion, hardware failure, or cyber-attacks. These backups will be securely stored and tested frequently to ensure their reliability and that they can be restored promptly when needed. By establishing a robust backup and recovery strategy, critical data can be quickly recovered in the event of an unexpected outage or attack.

By implementing these measures, both the VinDr-SpineXR dataset and the synthetic data created through the C-arm simulation can be protected and remain compliant with GDPR regulations, ensuring their availability when needed.

E. Data curation

Ensuring the accuracy and completeness of the VinDr-SpineXR dataset is crucial for any research project that utilizes it. To achieve this, a rigorous data curation process should be implemented that involves standardization of data formats, verification of data completeness, and removal of incomplete or inconsistent data points. Quality control measures will also be established to guarantee that the dataset meets the necessary standards for research purposes. These measures may include using standardized procedures for data collection, establishing quality control metrics, and conducting independent review and validation of data by multiple experts. Overall, implementing such measures will ensure that the VinDr-SpineXR dataset is of high quality and suitable for use in research projects that require reliable and accurate data.

V. Business Case

The proposed start-up will provide AI software for C-arm machines and has a significant business case due to the software's ability to reduce radiation dosages for patients and especially physicians. Radiation reduction for physicians is a significant advantage that can benefit both physicians and hospitals, specifically hospital administrators.

The software's robustness for different C-arm machines makes it highly scalable and marketable. The software will be compatible with different types of C-arm machines and can be sold to C-arm manufacturers or end-users such as hospitals and diagnostic centers, increasing its potential market reach. The software's scalability also provides an opportunity for the start-up to enter different healthcare markets, increasing its potential for growth.

A. Back of the envelope calculation

To estimate the market size and revenue for the start-up, a back-of-the-envelope calculation is performed using the TAM, SAM, and SOM analysis [27]. According to various market research reports [28] [29] [30], the global C-arm market size was valued at around USD 2 billion in 2021 and is expected to grow at a Compound Annual Growth Rate (CAGR) of 4.4% to 5.6% from 2022 to 2030, reaching USD 3 billion by 2030. The U.S. C-arm market size was USD 814 million in 2019 and is projected to reach USD 1.15 billion by 2027, exhibiting a CAGR of 4.4% during the forecast period [30]. The mobile C-arm segment accounted for more than 85% of the global and U.S. market share in 2021 [28] [30]. The major end-users of C-arm machines are hospitals, diagnostic centers, and specialty clinics [28] [29] [30].

We will focus on the fixed C-arm market, as this is more favorable when installing the hardware for edge computing. Furthermore, we will focus on the US market. Based on these data, we can assume that the TAM for the proposed start-up is 450 million USD by 2030, which is the global C-arm market size. There's no data available on the percentage of

C-arm market that is related to software, especially AI-related software. However, an estimation can be made by looking at other devices such as autonomous mobile robots. According to a report [31], the AI software market for autonomous mobile robots was estimated to account for approximately 61% of the total market for autonomous mobile robots in 2020. Taking into account that the adoption of AI in medical devices is slower than in other markets due to regulatory challenges for example, it is assumed that 25% of the global C-arm market accounts for AI software. Thus, a Serviceable Available Market (SAM) of 113 million USD is estimated. Focusing on the US market, we assume that 10% of this SAM can be realistically captured by the start-up in a given time frame, we can estimate the SOM as 11.3 million USD by 2030. These numbers indicate that there is a significant market opportunity for the proposed start-up and its AI software for C-arm machines. However, there are also some challenges and risks that need to be addressed, such as regulatory approval, clinical validation, customer adoption, and competition from other players in this field.

VI. Assumption mapping

Assumption mapping is a process that involves identifying and analyzing the assumptions that underpin a business model or product. In the context of the proposed start-up, assumption mapping can help to identify and test the assumptions made about the product, market, and customers. The assumptions are discussed the following subsection.

A. Assumptions

Assuming the algorithm can successfully reduce the noise in X-ray images without compromising the image quality or diagnostic accuracy, it should be able to handle different types of noise, such as quantum noise, that may arise from low-dose X-ray exposure. Additionally, the algorithm should be robust to variations in image acquisition parameters, such as exposure time, tube voltage, and tube current.

To integrate the software product with existing C-arm machines, it should be possible to install hardware. Edge computing will be used as described in section III-G. Compatibility with the C-arm machine is assumed, and the software should comply with the relevant standards and regulations for medical devices and data security. Moreover, it should not interfere with the normal operation of the C-arm machine or cause any delays or errors in image processing or transmission.

Demand for the software product among physicians who use C-arm machines for various procedures, such as angiography, fluoroscopy, and orthopedics, is assumed. The software product should offer clear benefits to physicians, such as improved image quality, reduced radiation exposure for patients and staff, and potentially resulting in enhanced diagnostic confidence. Additionally, the software product should be easy to use and cost-effective.

It is also assumed that there is a market opportunity for the software product among C-arm manufacturers and distributors. The software product should provide a value proposition to C-arm companies, such as increased sales, customer satisfaction, and brand reputation. Furthermore, it should be compatible with different models and brands of C-arm machines.

Another assumption is that the denoised X-ray images should have a similar or comparable contrast, resolution, and signal-to-noise ratio as normal X-ray images from the C-arm machine. Exploiting the sparsity property of X-ray images, which means that most of the pixels in an image have zero or near-zero values, is a fundamental assumption. This property allows for efficient compression and reconstruction of X-ray images. Additionally, it enables the removal of noises such as quantum noise, which are random fluctuations in pixel values due to low photon counts. White noise is one of the main sources of noise in low-dose X-ray imaging, and it can degrade image quality and diagnostic performance. It is assumed that the AI algorithm should be able to remove white noise and distinguish quantum noise from other types of noise or signal components and remove it accordingly.

The last assumption is that the software product should demonstrate superior performance, reliability, usability, or affordability compared to other AI products in terms of radiation reduction. It is a crucial assumption that the software product has a competitive advantage.

The assumptions can be tested using various methods, such as customer surveys, market research, prototype testing, and financial modeling. Testing the assumptions allows for validation or invalidation, and adjustments to the business model can be made accordingly. Assumption mapping is an ongoing process that should be revisited regularly as the business evolves and new information becomes available. Continual refinement of the understanding of the assumptions that underpin the business model can increase the likelihood of success and mitigate risks.

B. Methods to validate/invalidate assumptions

In this section, we describe the methods we used to test our major assumptions and evaluate the feasibility of our proposed solution. We focused on three key assumptions that are critical for the success of our project: the existence of a market demand, the effectiveness of our denoising algorithm, and the latency of our denoising algorithm. These three key assumptions are then tested with micro experiments as described below.

The first assumption we made was that there is a market demand for our product. To validate this assumption, we built a website that showcases our product and its features. As part of our promotional strategy, we reached out to potential customers through targeted email campaigns. Using carefully curated contact lists, we sent personalized emails to individuals who we believed would be interested in our product. These emails contained information about our product and its unique features, along with a direct link to our website. To track the effectiveness of our email promotions, we used analytic tools to monitor the traffic generated from these email campaigns. By analyzing the website traffic data, we gained insights into the level of interest and engagement from the recipients of our emails. Additionally, our website includes a contact form where visitors can provide their information, including their name, email address, and any messages they may have. This allowed us to gather feedback directly from potential customers and further validate their interest in our product. To consider our assumption validated, we have set a benchmark of receiving at least 20 inquiries or messages expressing genuine interest in our product. This significant number of interactions serves as a strong indication that our product is desired by potential customers. By leveraging targeted email promotions and closely monitoring website traffic and interactions, we aim to reach this milestone and gather valuable insights into the level of interest from potential customers.

The second assumption we made was that we can successfully reduce the noise in X-ray images without compromising the image quality or diagnostic accuracy. To validate this assumption, we applied our denoising algorithm to the VinDr-SpineXR data set. We compared the original and denoised images qualitatively using visual inspection and quantitatively using metrics such as peak signal-to-noise ratio (PSNR) and mean squared error (MSE). We considered this assumption to be validated if the PSNR is greater than 50 dB. As a later step and final validation, physicians can be asked to evaluate the images and rate them on a scale of 1 to 5 based on their quality and diagnostic usefulness. The assumption is validated if the average score of the physicians is equal or greater than 3.

The third assumption we made was that the latency of our system is low enough to provide a fast and smooth user experience. To validate this assumption, we measure the inference speed of our denoising algorithm. We considered this assumption to be validated if the inference speed of our algorithm was less than 300 ms on average.

VII. RESULTS

In this section, we provide a qualitative and quantitative analysis of the results, and discuss the implications and limitations of our approach. Furthermore, results of our micro experiments will also be discussed.

The qualitative analysis was based on visual inspection of the denoised images. A sample image and its corresponding denoised output are presented in fig. 4. The input image was corrupted with a high level of gaussian noise to simulate a challenging scenario that exceeds the expected noise level in realistic applications.

The figure illustrates the effectiveness of the model in denoising the noisy image. However, some limitations are also evident. The model tends to smooth out the finer details of the spine, especially those with thin edges. This is partly due to the high level of noise in the noisy image, which obscures the information about fine details. Nonetheless, the denoised image preserves the relevant information for clinical purposes. For instance, the physician can still distinguish between the different regions of the spine and locate the appropriate spot for needle insertion. To further evaluate the performance of the model, we conducted a quantitative analysis on the test dataset. The quantitative analysis was based on two metrics: peak signal to noise ratio (PSNR) and mean squared error (MSE). The scores obtained by the model can be seen in table I. Higher PSNR and lower MSE are better.

TABLE I: Scores of the model evaluated on the test set (VinDr-SpineXR dataset), Higher PSNR and lower MSE are better.

PSNR [dB]	**MSE [-]**
82.25	0.0004

The model achieved a high PSNR of 82.25 dB and a low MSE of 0.0004, which demonstrate its effectiveness in restoring high-quality images from low-quality inputs. These scores are comparable to or better than those reported by previous state-of-the-art methods on similar datasets [32]–[34]. The quantitative analysis of the results validates the second assumption as mentioned in section VI-B.

Since the denoising is intended for real-time use on the c-arm, the model has also been used to inference of samples that are correlated in the temporal domain, i.e. consecutive frames. Experiments have been conducted to improve the performance of the model on the frames of a video. Correspondingly, since the images are in grayscale, they can be stacked along the input channel dimension, which is 3 for RGB images. This allows the temporal information to be utilized to further enhance the denoising process. This technique significantly improves denoising results, as observed through visual inspection. It improves the preservation of edges and details, while significantly reducing flickering in the video [1].

The average inference time for the samples is 10.33 seconds, which is quite high. The high inference time can be fixed by altering the model or by increasing the computational power of the hardware. Due to limited time resources, no empirical results could be provided to support that claim. The high inference time means that the third assumption as mentioned in section VI-B has been invalidated.

Furthermore, as part of our broader market research, we conducted a microexperiment to validate the assumption of market demand for our product. By creating a website showcasing our product and using targeted email campaigns, we observed some initial traffic indicating early interest. However, due to the recent creation of the website, we did not have sufficient time to reach our benchmark of receiving at least 20 inquiries or messages expressing genuine interest. Nevertheless, based on the current results, we remain optimistic that with continued efforts and longer exposure time, we can attain the desired goal and gather valuable insights into the level of interest from potential customers, further validating the demand for our product. These findings from the microexperiment support the assumption that there is interest in our product, and provide a foundation for further investigation and market analysis.

[1]These results can be seen at: https://github.com/hakdemir/Denoising

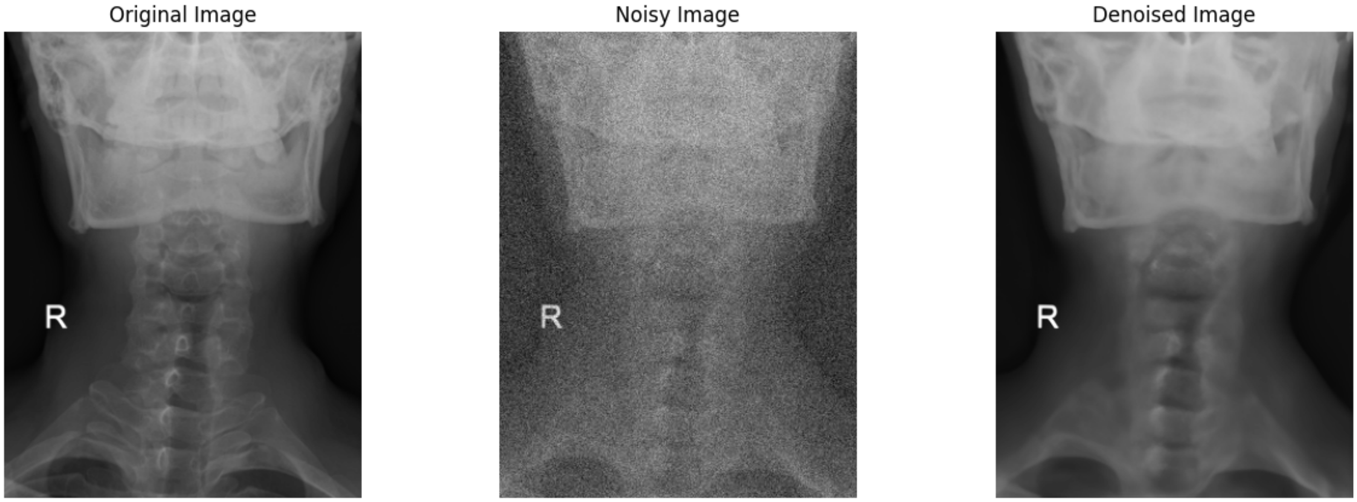

Fig. 4: Denoising results of a sample

VIII. Conclusion

The development of a deep learning algorithm for real-time denoising of low-dose X-ray images in C-arms will be a compelling business case for the potential start-up. The proposed methodology is thoroughly researched and tested through six maturity steps. Various techniques for mitigating the effects of lowering radiation dosages, including deep learning-based methods such as convolutional neural networks (CNNs), autoencoders, and transformers. The performance of denoising models is evaluated using metrics such as peak signal-to-noise ratio (PSNR) and mean squared error (MSE). A KBNet baseline was chosen to do the denoising task due to its ability to implement state-of-the-art methods while keeping computational cost relatively low. Six maturity steps have been defined. These steps involve evaluating the denoising model with various datasets, including CIFAR-10 and VinDr-SpineXR, as well as simulated data generated using a C-arm simulator. Data de-identification, storage, management, curation, and acquisition are also considered for the data. The business case section presents a strong argument for the proposed start-up, which will provide AI software for C-arm machines. The software's ability to reduce radiation dosages for patients and especially physicians is a significant advantage that can benefit both hospitals and physicians by reducing costs and minimizing health risks. The software's robustness for different C-arm machines makes it highly scalable and marketable. A back-of-the-envelope calculation using TAM, SAM, and SOM analysis estimates a significant market size and revenue potential for the start-up. This innovation has the potential to improve medical imaging procedures, advance medical diagnosis and treatment, and reduce healthcare costs.

While our experimental results using our current image denoising model with pretrained weights have shown promising outcomes, we believe there is untapped potential for further improvement through the utilization of transfer learning techniques [35]. Transfer learning involves leveraging knowledge learned from one task (pretraining) and applying it to another related task (denoising) by fine-tuning the model. By retraining the network using our specific dataset, we hypothesize that the model can learn to adapt its learned features to better handle the noise characteristics present in our data. This process involves updating the weights of the model's last few layers while keeping the majority of the network's parameters frozen. We anticipate that by fine-tuning the pretrained model on our dataset, we can improve its performance in noise reduction, preserving important image details, and handling diverse noise patterns commonly encountered in real-world scenarios. Additionally, the transfer learning approach may increase the model's robustness and generalization capabilities, enabling it to effectively denoise images with varying noise levels.

In addition to exploring transfer learning, another avenue for improving the performance of the image denoising model is by adapting it to consider the temporal domain. By incorporating temporal information, such as the temporal coherence between consecutive frames in a video sequence, we can potentially enhance the denoising process and produce even cleaner and more visually pleasing results. The temporal is already taken into account when the image sequence is denoised using stacked images. However, there are other, more effective approaches. One approach to incorporate temporal information is by utilizing 3D convolutions [36]. Traditional 2D convolutions operate on individual frames independently, while 3D convolutions extend this concept by considering a sequence of frames as a 3D volume. This allows the network to capture temporal dependencies and exploit temporal coherence in the denoising process. By incorporating 3D convolutions into the architecture, the model can effectively learn spatio-temporal features that are crucial for denoising video sequences. The additional dimension enables the network to capture motion information and exploit temporal correlations, resulting in improved denoising performance and reduced temporal artifacts. However, it is important to note that incorporating 3D convolutions introduces additional computational complexity and memory requirements. Training a network with 3D convolutions may also require a larger dataset of video sequences to capture a diverse range of temporal patterns effectively.

In addition to focusing on the denoising performance, it is crucial to consider the inference time of the model, as it directly impacts the practicality and real-time applicability of the denoising system. Currently, the inference time is relatively high, which may hinder its deployment in time-sensitive scenarios. To address this challenge, two primary strategies can be explored: increasing computational power or decreasing the complexity of the network. Increasing computational power, such as utilizing high-performance hardware or distributed computing, can potentially reduce the inference time by leveraging parallel processing capabilities. However, this solution might come at an increased cost and resource requirements. Alternatively, reducing the complexity of the network can also lead to improved inference time. This can be achieved through architectural modifications, such as reducing the number of layers, employing lightweight modules, or utilizing efficient network designs. By optimizing the network structure and parameters, we can strike a balance between denoising performance and inference time.

References

[1] World Health Organisation, "Ionizing radiation, health effects and protective measures," 2016, https://www.who.int/news-room/fact-sheets/detail/ionizing-radiation-health-effects-and-protective-measures, Last accessed on 2023-03-30.

[2] The U.S. Department of Energy Office of Environment, Health, Safety Security , "Radiation in perspective," 2016, https://www.energy.gov/sites/prod/files/2016/09/f33/Radiation_in_Perspective.pdf, Last accessed on 2023-03-29.

[3] EUROPEAN COMMISSION Unit D.4 — Radiation Protection , "Radiation protection no 166," 2010, https://energy.ec.europa.eu/system/files/2014-11/166_0.pdf, Last accessed on 2023-03-29.

[4] C. H. Jung, J. S. Ryu, S. W. Baek, J. H. Oh, N. S. Woo, H. K. Kim, and J. H. Kim, "Radiation exposure of the hand and chest during c-arm fluoroscopy-guided procedures," *The Korean journal of pain*, vol. 26, no. 1, pp. 51–56, 2013.

[5] C. Tian, L. Fei, W. Zheng, Y. Xu, W. Zuo, and C.-W. Lin, "Deep learning on image denoising: An overview," *Neural Networks*, vol. 131, pp. 251–275, 2020.

[6] W. Jifara, F. Jiang, S. Rho, M. Cheng, and S. Liu, "Medical image denoising using convolutional neural network: a residual learning approach," *The Journal of Supercomputing*, vol. 75, pp. 704–718, 2019.

[7] M. Gholizadeh-Ansari, J. Alirezaie, and P. Babyn, "Low-dose ct denoising with dilated residual network," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2018, pp. 5117–5120.

[8] C. Tian, Y. Xu, Z. Li, W. Zuo, L. Fei, and H. Liu, "Attention-guided cnn for image denoising," *Neural Networks*, vol. 124, pp. 117–129, 2020.

[9] Koninklijke Philips N.V., "Clinically proven azurion with clarityiq technology," https://www.philips.com.au/healthcare/resources/landing/azurion/articles/clarityiq, Last accessed on 2023-03-30.

[10] Siemens Medical Solutions USA, Inc., "Ai rad companion, ai-powered decision support in chest x-ray, note = https://marketing.webassets.siemens-healthineers.com/208f1cb920b1140e/50d20396f8b5/siemens-healthineers-da_ai-rad_companion_chest_x-ray_editorial_whitepaper_ridder.pdf, last accessed on 2023-03-30,."

[11] Canon Medical Systems USA, Inc., "Celex - cxdi-rf," https://eu.medical.canon/products/xray/fluoroscopy/celex/technology, Last accessed on 2023-03-30.

[12] GE HealthCare, "Critical care suite," https://www.gehealthcare.com/products/radiography/mobile-xray-systems/critical-care-suite-on-optima-xr240amx, Last accessed on 2023-03-30.

[13] E. Samei and M. J. Flynn, "Quantitative imaging: radiomics, dosimetry, and personalized medicine," *Academic radiology*, vol. 19, no. 7, pp. 805–816, 2012.

[14] S. Singh, M. K. Kalra, S. Rathnayake, J.-A. O. Shepard, and S. R. Digumarthy, "Effect of radiation dose reduction on image quality and noise in chest radiography," *Journal of Thoracic Imaging*, vol. 33, no. 1, pp. 53–59, 2018.

[15] T. M. Quan, T. Nguyen-Duc, M. J. Kim, and G.-H. Jeon, "X-ray ct image reconstruction using a sparse-view deep learning network," *Journal of X-Ray Science and Technology*, vol. 23, no. 3, pp. 327–339, 2015.

[16] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*. Springer, 2015, pp. 234–241.

[17] D. Wang, F. Fan, Z. Wu, R. Liu, F. Wang, and H. Yu, "Ctformer: convolution-free token2token dilated vision transformer for low-dose ct denoising," *Physics in Medicine & Biology*, vol. 68, no. 6, p. 065012, 2023.

[18] Y. Zhang, D. Li, X. Shi, D. He, K. Song, X. Wang, H. Qin, and H. Li, "Kbnet: Kernel basis network for image restoration," 2023.

[19] A. Abdelhamed, S. Lin, and M. S. Brown, "A high-quality denoising dataset for smartphone cameras," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

[20] IBM, "The edge computing continuum: A vision for cloud-edge synergy," 2017, https://www.research.ibm.com/articles/edge-computing.shtml, last accessed on 2023-05-16.

[21] A. Krizhevsky, "Learning multiple layers of features from tiny images," University of Toronto, Tech. Rep., 2009. [Online]. Available: https://www.cs.toronto.edu/~kriz/learning-features-2009-TR.pdf

[22] H. H. Pham, H. Nguyen Trung, and H. Q. Nguyen, "Vindr-spinexr: A large annotated medical image dataset for spinal lesions detection and classification from radiographs (version 1.0.0)," https://physionet.org/content/vindr-spinexr/1.0.0/, 2021.

[23] A. Krizhevsky, G. Hinton *et al.*, "Learning multiple layers of features from tiny images," 2009.

[24] K. T. Nguyen, T. H. Khuat, T. A. Nguyen, N. T. Nguyen, T. V. Nguyen, and H. Van Nguyen, "A novel deep learning approach for multi-task classification of spinal disorders," *Computer Methods and Programs in Biomedicine*, vol. 201, p. 105918, 2021.

[25] T. Duong, T. Le, and T. Pham, "Spinenet: Automated classification and localization of common spine diseases using a deep neural network," *Journal of Medical Systems*, vol. 45, no. 2, pp. 1–12, 2021.

[26] T. A. Ngo, K. T. Nguyen, H. V. Nguyen, T. V. Nguyen, and T. A. Nguyen, "Automated detection and classification of spinal diseases using a convolutional neural network," *Computerized Medical Imaging and Graphics*, vol. 90, p. 101888, 2021.

[27] Similarweb, "Market sizing: How to calculate your total addressable market," 2023, https://www.similarweb.com/blog/research/market-research/market-sizing/, last accessed on 2023-05-16.

[28] "C-arms market size, share covid-19 impact analysis, by type (fixed and mobile), by application (orthopedics and trauma, cardiology, neurology, gastroenterology, oncology, and others), by detector (flat panel detectors and image intensifier), by end-user (hospitals, diagnostic centers, specialty clinics, and others), and regional forecast, 2022-2029," 2023.

[29] "Fixed and mobile c-arms market size, share trends analysis report by product (fixed, mobile), by application, by region, and segment forecasts, 2022 - 2030," 2023.

[30] "U.s. c-arms market size, share covid-19 impact analysis, by type (fixed c-arms, mobile c-arms (full-size c-arms, mini c-arms), by application (orthopedics and trauma, cardiology, neurology, gastroenterology, oncology, and others), by detector (flat panel detector, and image intensifier), by end user (hospitals, ambulatory surgical center, and specialty clin-ics), 2020-2027," 2020.

[31] "Autonomous mobile robot market by type, by application, by end-user - global opportunity analysis and industry forecast 2022-2030," Research and Markets, Tech. Rep., 2022. [Online]. Available: https://www.grandviewresearch.com/industry-analysis/industrial-robotics-market/segmentation

[32] M. A. Baig, A. A. Moinuddin, and E. Khan, "Psnr of highest distortion region: an effective image quality assessment method," in *2019 Interna-*

tional Conference on Electrical, Electronics and Computer Engineering (UPCON). IEEE, 2019, pp. 1–4.

[33] M. Patwari, R. Gutjahr, R. Raupach, and A. Maier, "Limited parameter denoising for low-dose x-ray computed tomography using deep reinforcement learning," *Medical Physics*, 2022.

[34] Z. Liu, Y. Zhang, and S. Wang, "Chest x-ray image denoising method based on deep convolution neural network," *IET Image Processing*, vol. 13, no. 14, pp. 2737–2744, 2019.

[35] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, and Q. He, "A comprehensive survey on transfer learning," *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43–76, 2020.

[36] Y. Cai, L. Ge, J. Liu, J. Cai, T.-J. Cham, J. Yuan, and N. M. Thalmann, "Exploiting spatial-temporal relationships for 3d pose estimation via graph convolutional networks," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 2272–2281.