Deep Monochromatic Metal Artifact Reduction for Computed Tomography

Registration Number: Comp-066

Student:

Sally Sijie Song

The Independent Schools Foundation Academy, Hong Kong

Supervising Teacher:

Dr. Simon D.J. Griffin

Secondary Teacher and Shuyuan Research Scientist
The Independent Schools Foundation Academy, Hong Kong

August 31, 2021

Abstract

Computed tomography (CT) is a 3-dimensional medical imaging modality that uses X-ray beams to generate cross-sectional images of the human body. Due to its ability to accurately depict the internal structure of the body, CT is widely used in medical diagnosis. However, the diagnostic value of CT images is often severely impaired by metal artifacts, which are errors in the image induced by the presence of metal implants. While most existing metal artifact reduction (MAR) algorithms operate on data acquired by a single X-ray energy spectrum, this study proposes a model that operates on data generated by dual-energy CT, a less popular CT technique that acquires projection information using 2 distinct energy spectra to achieve a higher capability for material differentiation. The proposed MAR model is a novel end-to-end convolutional autoencoder-based network that achieved outstanding performance (Structural Similarity Index > 0.994, Mean Squared Error < 7×10^{-6}, Peak Signal-to-Noise Ratio > 58 dB) on a remarkably small training set of 390 instances. Since MAR networks are commonly trained on manually synthesized metal artifact simulations, this study also proposes a data synthesis method that replaces the manually segmented realistic metal implant shapes used in conventional metal artifact data simulation with randomly generated polygons. The proposed data synthesis method vastly improves the efficiency of the data generation process.

2

Contents

1 Introduction

1.1 CT Imaging

Computed tomography (CT) is a non-invasive imaging technique that represents a major advance for medical diagnostics. It provides comprehensive 3-dimensional information with spatial and contrast resolution far exceeding that possible in conventional 2-dimensional X-ray images.

A CT scan is conducted by aiming an X-ray beam at the patient while rotating the beam around the body. After penetrating the body, the X-ray beam is received by CT detectors that generate signals that are then processed by the computer to form projections. All of the projections taken at different projection angles are stacked together horizontally to create a sinogram, with each horizontal line on the sinogram representing a projection taken at a single projection angle. After data acquisition, the sinogram is reconstructed into a readable image using reconstruction algorithms. Reconstructed images can then be interpreted by a physician [42]. The most commonly used reconstruction algorithm is the filtered back projection (FBP) algorithm, which this study uses to simulate reconstructed images.

The attenuation coefficient quantifies the capability of a material to absorb and weaken (attenuate) an incident energy beam [55]. CT uses the attenuation of the X-ray beam to distinguish between different materials: for example, materials with lower attenuation coefficients, such as soft tissues, tend to attenuate the beam less than materials with higher attenuation coefficients, such as bones and metals (see Figure 1).

In CT images, attenuation is reflected in the voxel value (a voxel being the 3-dimensional equivalent of a pixel). When produced as a result of a beam's passage through adjacent regions of the body that differ in their degree of attenuation, a discernible difference in voxel values defines the contrast resolution of the CT image.

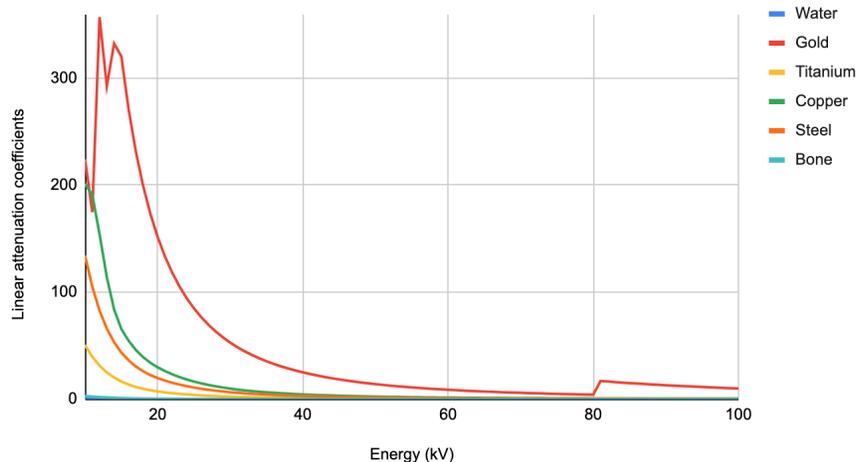Linear attenuation coefficients of materials vs. X-ray energy levels

Figure 1: Attenuation coefficient of different materials as a function of energy [46]. The attenuation coefficient of a material is generally correlated with its density.

1.2 Metal Artifacts in CT

The clinical plausibility of a CT image is dependent on the image's fidelity to the scanned object. Therefore, the suppression of CT artifacts is often essential for the clinician's diagnostic confidence. CT artifacts are errors in the CT image. They are especially severe when induced by the presence of metal implants. Metal implant-induced artifacts, usually referred to as metal artifacts, can deteriorate CT image quality to non-diagnostic degrees.

Most metal artifacts are caused by beam hardening, a process that occurs due to the reconstruction algorithm's inability to accommodate for the polychromatic nature of the X-ray spectrum. To aid understanding, a polychromatic X-ray spectrum is visualized in Figure 2. When an X-ray beam penetrates an object, the lower-energy photons in the X-ray are more easily attenuated than the higher-energy photons. The polychromatic beam, therefore, loses the lower-energy part of its spectrum, which causes the mean intensity of the transmitted X-ray beam to increase, or "harden", and become higher than the mean intensity of the initial X-ray beam [1].
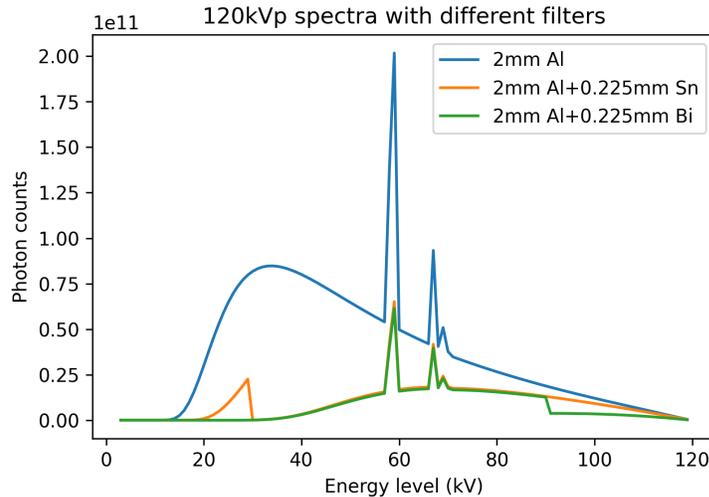
Figure 2: The X-ray spectrum at the 120kV tube voltage when different filters are used. Filters (metal sheets placed between the X-ray beam and the patient) are commonly used to filter out some low-energy photons to partially reduce metal artifacts.

Beam hardening produces streaks around high-density objects, such as bones or metal implants in a clinical CT image (see Figure 3). After a beam passes through a hard material (a material with a high attenuation coefficient) and becomes "hardened", it is attenuated much less when it then passes through soft materials as the lower-energy photons in the beam have already been lost. This results in a lower voxel value of the softer material, making it appear darkened on the image [1]. Streaks can obscure key information on the image and impede medical evaluation.

(a) Streaks

(b) Artifact-free image

Figure 3: The effect of streaks modeled by 3 plastic cylinders (soft material) surrounded by 3 metal cylinders (hard material). Dark and bright streaks between the metal cylinders impede the inspection of the plastic cylinders. [6]
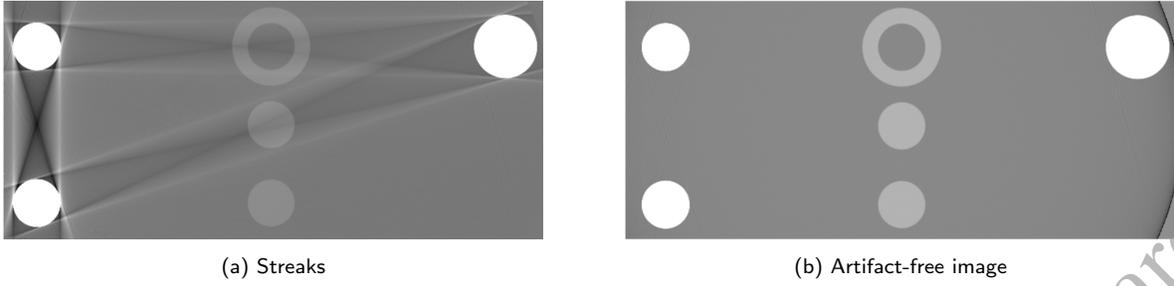
Beam hardening only occurs in CT machines that use energy-integrating detectors (EIDs). EIDs calculate signals by integrating the total energy deposited by all transmitted photons. On the contrary, photon-counting detectors (PCDs) can register the energies of individual photons to attain spectral information and avoid beam hardening. However, due to the high costs of PCDs, they are not typically clinically available [43]. Most CT images, therefore, suffer from beam hardening artifacts.

1.3 Theory

Because beam hardening is the main source of metal artifacts, this study aims to perform MAR through reducing beam hardening-induced artifacts. In this section, the mathematical theory behind the beam hardening reduction task is introduced in detail.

1.3.1 The Main Source of Metal Artifacts: Beam Hardening

This subsection explains the mathematics behind beam hardening.

Diagnostic CT imaging devices use polychromatic X-ray beams to generate attenuation information. When a polychromatic X-ray beam penetrates a single material, the relationship between the incident (initial) beam intensity I_0 and the transmitted X-ray beam intensity I can be modeled by the following logarithmic equation:

$$
\begin{aligned}
\ln\left(\frac{I_0}{I}\right) &= P_p \\
&= \ln\left(\frac{\int S(E)dE}{\int S(E)e^{-\int_0^l \mu(x,E)dx}dE}\right)
\end{aligned}
\tag{1}
$$

where P_p denotes the logarithmic polychromatic projection value, $\mu(x,E)$ denotes the energy-dependent attenuation coefficient at position x along the ray under the energy level of E, and $S(E)$ denotes the spectrum of X-ray energies.

However, the polychromatic projection value must be reconstructed to form a readable monochromatic image. After the polychromatic projection sinogram is first pre-processed with a mathematical process involving the Fourier transform, it is then reconstructed with the reconstruction algorithm to form a monochromatic image [61]. This study simulated beam hardening effects with the most commonly used reconstruction algorithm: the FBP algorithm

6

[42]. According to the FBP, the relationship between I_0 and I after the X-ray penetrates a single material can be modeled by the following equation:

$$-\ln\left(\frac{I}{I_0}\right) = P_m = \int_0^l \mu(x)dx \tag{2}$$

where P_m denotes the logarithmic monochromatic projection value and $\mu(x)$ denotes the attenuation coefficient at position x along the ray.

As shown in Equation 2, the FBP makes 2 assumptions: 1) the X-ray spectrum is monochromatic and 2) the attenuation coefficient is energy-independent. Because the assumptions made by the FBP overlooks both the spectral nature of the X-ray beam and the energy dependence of the attenuation coefficient, the FBP reconstruction introduces discrepancies between the physical reality of the object and the reconstructed projection of the object. These discrepancies produce beam hardening artifacts in the reconstructed image.

1.3.2   Mathematics of Beam Hardening Correction

Since beam hardening is caused by the disparity between the polychromatic projection value P_p and the reconstructed monochromatic projection value P_m, the goal of beam hardening correction is to minimize the difference between the P_p and P_m of each voxel. This minimization can be done through linearization, which aims to predict the P_m value through P_p data. Single-material linearization was proposed by Herman and Trivedi [44]. Their linearization method suggests that when the X-ray beam penetrates a single material, the relationship between P_p and its corresponding monochromatic value P_m can be modeled by the following polynomial:

$$P_m = f(P_p) = a_n P_p^n + a_{n-1} P_p^{n-1} + ... + a_0 \tag{3}$$

where n is the order of the polynomial and $a_0, a_1, a_2, \ldots, a_n$ are the coefficients of the polynomial. Typically, n is a value between 1 and 4, but higher polynomial degrees may be needed for the linearization of high-density materials.

Since the single-material linearization method introduced above can only perform beam hardening correction on objects with homogeneous material compositions, Peter M. Joseph and Robin D. Spital [45] proposed a multi-material linearization method to be used for objects with heterogeneous material compositions. Specifically, this method assumes that the object is comprised of 2 materials: a hard material with a higher attenuation coefficient and a soft material with a lower attenuation coefficient. The relationship between the I_0 and I after passing through the object is therefore modeled as:

$$-\ln\left(\frac{I}{I_0}\right) = p_m = \int \mu_H dl_H + \int \mu_S dl_S \tag{4}$$

where p_m denotes the logarithmic monochromatic projection value, μ_H and μ_S denote the attenuation coefficients of the hard and soft materials respectively, l_H and l_S represent the path lengths (distance traveled by the X-ray in the material) of the hard and soft materials respectively.

The polychromatic projection value can be modeled as:

$$p_p = -\ln\left(\frac{\int S(E)e^{-\int_0^{l_H}\mu_1(x,E)dx-\int_0^{l_S}\mu_2(x,E)dx}dE}{\int S(E)dE}\right)$$

$$= -\ln\left(\frac{\int S(E)e^{-\int_0^{l_H}\mu_1(x,E)dx}dE}{\int S(E)dE}\right) - \ln\left(\frac{\int S'(E)e^{-\int_0^{l_S}\mu_2(x,E)dx}dE}{\int S'(E)dE}\right) \tag{5}$$

$$= p_H + p_S$$

where $S'(E)$ denotes the spectrum that has been "hardened" by the hard material that has propagation length l_H, p_p denotes the logarithmic polychromatic projection value, and p_H and p_S denote respectively the polychromatic contributions of the hard and soft materials.

The relationship between $S'(E)$ and $S(E)$ can be represented by:

$$S'(E) = S(E)e^{-\int_0^{l_1}\mu_1(x,E)dx} \tag{6}$$

There are 2 stages of multi-material linearization. The first stage involves building a lookup table (LUT) for soft material correction. The propagation length of the hard material is then discretized and each discrete propagation length l_H is assigned to a certain step. For each step, a polynomial fit f is used to linearize the polychromatic contribution of the soft material (p_S) to the monochromatic contribution (m_S) under the corresponding $S'(E)$ using Equation 7.

$$m_S = f(p_s) = a_n p_s^n + a_{n-1}p_s^{n-1} + ... + a_0 \tag{7}$$

where n is the order of the polynomial and $a_0, a_1, a_2, \ldots, a_n$ are the coefficients of the polynomial.

The second stage is the correction stage which takes place after the initial reconstruction. During this stage, the hard material is first segmented. Then, l_H is used to index from the LUT to correct the soft material. The monochromatic contribution of the hard material is calculated with the following equation:

$$m_H = \mu_H(E_0) \times l_H \tag{8}$$

where E_0 is the monochromatic energy level.

1.4   Related Works

Linearization is the theoretical basis for beam hardening correction-focused MAR methods in that the goal of MAR is to transform the polychromatic projection into a monochromatic projection. Instead of using the polynomial $f(x)$ in Equation 3 to achieve this goal, however, existing MAR methods implement algorithms to correct beam hardening.

Although MAR has been a popular field of research and development for 4 decades [15], there is still no universal solution to the problem [16], [17]. Existing literature has explored both conventional interpolation-based and deep learning-based MAR methods.

1.4.1 Conventional MAR

Conventional MAR algorithms can be roughly classified into 3 groups — 1) pre-processing, 2) post-processing, and 3) dual-energy CT (DECT). The drawbacks of the most popular conventional MAR algorithms are summarized in Table 1

Pre-processing methods are applied to the sinogram. They attempt to model the physical source of metal artifacts (*e.g.* beam hardening [56], [57], [58] or photon starvation [59]) to correct the raw data accordingly [10]. Pre-processing methods are not effective in removing artifacts caused by the presence of high-density metals [35].

Post-processing methods are applied to the reconstructed image. Popular post-processing methods include linear interpolation (LI) and iterative reconstruction. LI replaces the corrupted area of the projection with approximations made with neighboring uncorrupted pixels. Approximations made by LI tend to be accurate if *a priori* information about the body composition of the scanned area is incorporated in the forward projection [11], [12], [13], [14]. However, when *a priori* information is unavailable, the values estimated by LI are often incorrect [7]. The iterative reconstruction (IR) algorithm is used to generate images from uncorrupted projections [18], [19], [20], [21] or projections that have been corrected through other algorithms [22]. IR methods are usually effective, but they are computationally inefficient due to the large number of iterations required.

Unlike conventional single-energy CT that uses 1 X-ray spectrum to acquire attenuation information, DECT uses 2 distinct X-ray energy spectra that are usually produced at tube voltages of 80kVp and 140kVp [23], [24]. The difference in attenuation information acquired by the 2 spectra allows for better differentiation of the material composition [9]. The extra spectral information acquired increases the size of the dataset, which is the main reason that relatively fewer studies have focused on DECT-based MAR [51]. Both sinogram domain and reconstruction image domain DECT methods have been proposed, but the former normally works better than the latter because the images in the reconstruction domain have already been contaminated by metal artifacts.

Method name	Drawbacks	Works
Pre-processing	Ineffective in the presence of high density-metals	[56], [57], [58], [59]
LI	Requires information about the material composition of the scanned area	[11], [12], [13], [14], [7]
IR	Computationally inefficient	[18], [19], [20], [21], [22]
DECT	Requires large dataset size	[23], [24], [9]

Table 1: The drawbacks of different conventional MAR methods.

1.4.2 Deep Learning-Based MAR

To tackle the shortcomings of conventional MAR methods, deep learning-based methods were developed. The first deep learning-based MAR method was developed in 2017 by Gjesteby,

et al. [28]. Since then, there has been an increasing body of literature on deep learning-based MAR methods, most of which claim superior performance to conventional MAR methods. Within the existing works, both supervised and unsupervised methods have been proposed.

Supervised methods are trained using metal artifact-corrupted images as the input and metal artifact-free images as the target. The nature of supervised learning requires that every input image be paired with an anatomically identical target image. Because this requirement is impractical to fulfill within the clinical setting, most supervised methods resort to simulating paired metal artifacts-corrupted and uncorrupted CT images to generate the paired training set. Within supervised learning, the convolutional neural network (CNN) architecture has achieved success on both normal dose CT [35], [36] and low dose CT (a type of CT scan that provides 1.4 mSv of radiation instead of the 7mSv that regular CT provides) [30], [31], [32], [33], [34]. Most notably, Park et al. [29] employed a supervised U-Net [38] for MAR that achieved promising results on titanium hip prostheses. A unique CNN model that this study took inspiration from is the Butterfly-Net [62], which was not developed for MAR but was designed to decompose 2 base materials, bone and tissue, in metal implant-free DECT images. The Butterfly-Net operates on the image domain and does not employ encoding or decoding. Its main innovation is its double-entry double-out architecture that splits the model into 2 parallel material decomposition pathways, with each pathway responsible for the decomposition of 1 material. The model takes 2 dual-energy images as input and outputs the 2 decomposed material masks. An innovative feature of the Butterfly-Net is its implementation of a crossover architecture that allows information exchange between the 2 material generation pathways.

The main advantage of unsupervised learning methods is their ability to learn with unpaired data, which allows them to overcome the difficulty of acquiring paired metal artifact-corrupted and uncorrupted CT images. The training set of unsupervised learning methods usually consists of a small number of labeled images combined with a large number of unlabeled images. A commonly used unsupervised learning model is the Cycle-GAN model [40], [39], which is a generative adversarial network that employs a cycle consistency loss to allow for unpaired image learning. The Cycle-GAN used a U-Net [38] structure to perform upsampling and downsampling of the input image. Liao, et al. [37] proposed the artifact disentanglement network (ADN) that separates, or "disentangles" the metal artifact and the content component (anatomical structure) by separately processing the encoding and decoding of the artifact and content components for the unpaired inputs. The ADN is operated on synthesized data, which is easier to collect than real clinical data.

Lyu, Yuanyuan, et al. [41] sought to combine the advantages of both supervised and unsupervised MAR methods with the development of an unpaired dual-domain network (U-DuDoNet) trained using unpaired data and which operates on both the sinogram and the image domain. It directly simulates the process of artifact generation using two U-Nets, one on each domain.

2 Approach

2.1 Motivation

A novel end-to-end convolutional autoencoder-based MAR network is proposed. The network is trained on dual-energy data so as to utilize DECT's advantage over single-energy CT in material decomposition tasks [50]. Similar to the U-DuDoNet [41], the proposed model is designed to combine the advantages of both supervised methods and unsupervised methods.

Although CNNs are known for being powerful at image processing [52], the performance of a CNN is heavily reliant on the size of its training dataset [53]. The large dataset size typically required for CNN training not only reduces the CNN's computational efficiency but also increases the difficulty of data acquisition, especially when the model trains on clinical data whose availability is limited by patient privacy protection policies. To reduce the training dataset size and maximize the CNN's learning efficiency, data augmentation was performed to include a greater material diversity in the training dataset.

The proposed model, similar to most other deep learning-based MAR models, is trained on simulated metal artifact images created by overlaying metal implant-free CT images with realistic metal implant shapes. In an effort to reduce the extensive time and effort spent on the manual collection, segmentation, and pre-processing of realistic metal implant masks, an automatic synthetic implant generation method was developed to replace realistic metal implant shapes conventionally used for metal implant simulation with randomly generated polygons. At the core of the proposed CNN framework is an end-to-end auto-feature extraction process that minimizes the need for human supervision.

The innovations of the proposed approach are summarized below:

1. **A novel end-to-end convolutional autoencoder MAR network**

2. **A systematic data generation method that improves the performance and efficiency of the proposed network**

2.2 Data Simulation

The proposed model is a supervised learning method that requires the training set to consist of pairs of metal artifact-corrupted input images and uncorrupted yet otherwise identical target images. Due to the impracticality of acquiring such pairs of images in real clinical datasets, these pairs of images were simulated.

Metal implant-free CT images were overlaid with metal implant simulations to create uncorrupted virtual monochromatic target images (VMIs) that are used as images and metal artifact-corrupted input images.

2.2.1 Clinical Data for Simulation

Metal implant-free CT images with dimensions of 512×512 pixels consisting of bone and tissue were used as bases to lay metal implant simulations upon. These images were acquired from the open-source Low Dose CT Grand Challenge (LDCT) dataset and DeepLesion dataset [47] to simulate realistic metal artifact-corrupted images. From the LDCT dataset, 90 slices of normal dose CT scans were chosen: 30 slices of the hip area and 60 of the abdomen area. To

11

diversify the training set, the slices were taken from 5 different patients. From the DeepLesion dataset, 40 slices from a single patient were chosen. Of the slices, 4 are of the shoulder area, 20 are of the head area, 12 are of the teeth area, and 4 are of the hip area.

2.2.2 Simulation of Metal Artifact-Corrupted Sinograms

The general workflow of simulating metal artifact-corrupted sinograms involves the following steps: 1) thresholding the implant-free CT images slice by slice to calculate the material component of each voxel, 2) overlaying the implant-free images with the simulated metal implant masks, 3) ray tracing to create masks of each material, 4) calculating the metal artifact-corrupted polychromatic sinogram using the material masks. In step 4 of the workflow, 2 distinct polychromatic sinograms were created at the simulated tube voltages of 80 and 140kVp, respectively. The sinograms created at 80 and 140kVp are the dual-energy inputs to the MAR network.

The DECT spectra at the tube voltages of 80 and 140kVp were generated with the SPEKTR 3.0 tool [48]. Ray tracing of each material was performed in the ASTRA toolbox [49] using the Siddon algorithm. The reconstruction of the polychromatic sinogram was based on Equation 1. The attenuation coefficients of bone, soft tissue, and the metal of choice were included in the calculation process of the metal artifact-corrupted input sinogram. Since metal artifacts in the sinogram are created by the inclusion of the metal attenuation coefficient, the metal mask is not included in the calculation process of metal artifact-free target sinograms: only the bone and soft tissue masks are reconstructed.

The soft thresholding method used to assign voxels to different materials is in accordance with the method used by Zhang, Yanbo, et al. [35]. Voxel values (attenuation coefficients) below a certain threshold T_1 are assumed to contain soft tissue, whereas voxel values above a higher threshold T_2 are assumed to contain bone. Voxel values between T_1 and T_2 are assumed to be a mixture of the 2 materials. A weighting function for bone is introduced as,

$$w(x_i) = \begin{cases} 0 & x_i \leq T_1 \\ 1 & x_i \geq T_2 \\ \frac{x_i - T_1}{T_2 - T_1} & T_1 < x_i < T_2 \end{cases} \tag{9}$$

where x_i is the ith voxel value of x. Therefore, the soft tissue value x_s and bone value x_b can be expressed as,

$$x_i^b = w(x_i)x_i \tag{10}$$
$$x_i^w = (1 - w(x_i))x_i \tag{11}$$

2.2.3 Experimental Data Preparation

2 groups of metal implant masks were used for the simulation of corrupted and uncorrupted images. The first group of masks were realistic metal implant shapes conventionally used for the simulation of metal artifacts. The second group of masks consists of automatically generated random polygons. The training was performed separately on the 2 groups of artifact simulations to compare the performance of the models trained on the 2 data generation methods. 7 masks from each group were used for training. Each mask was assigned to multiple CT slices taken of 1 body part.

Data augmentation was performed on both groups of metal implant masks to include a large diversity of materials, shapes, and sizes. The diversifying of the training set improved the generalizability of the CNN so that it can operate on an extremely small training set of 390 instances, out of which 130 are corrupted input images simulated at the tube voltage of 80kVp, 130 are corrupted input images simulated at the tube voltage of 140kVp, and 130 and uncorrupted target virtual monochromatic images.

The dataset of realistic metal implant masks consists of 7 manually segmented metal implants. 4 common types of metal implants were simulated. They are 1) fixation screws, 2) hip prostheses, 3) dental fillings, and 4) surgical clips. The materials for the simulated implants are: titanium, steel, copper, gold. The attenuation coefficients of simulated materials were acquired from the NIST dataset [46]. Figure 4 shows all 7 realistic metal implants used for training. The images in the figure are cropped from the initial image dimensions of 512 × 512 pixels to better focus on the metal implants.

The general workflow of the pre-processing required for metal implant simulation is given in Figure 5. To simulate the fixation screws (Figure 4c, 4b, 4d, 7c, 7b), CAD drawings of commonly used fixation screw shapes were used as models. The models were rotated and resized in a Python environment using CV2 and PIL libraries to create a metal implant mask. The mask was then converted into a binary image for further use. Surgical clips (Figure 4e) and dental fillings (Figure 4f) were masked from the DeepLesion dataset, and the hip prostheses masks (Figures 4a, 7a) were manually drawn on the image on appropriate joint locations. To increase the shape diversity of the group 1 training set and to prevent overfitting, a gold piece (Figure 4g) with a randomly generated shape meant to resemble what a metal implant may realistically look like in certain images was added to the dataset.

(a) Titanium hip prostheses

(b) Copper long screw

(c) Titanium pedicle screws

(d) Steel pedicle screw

(e) Copper surgical clips
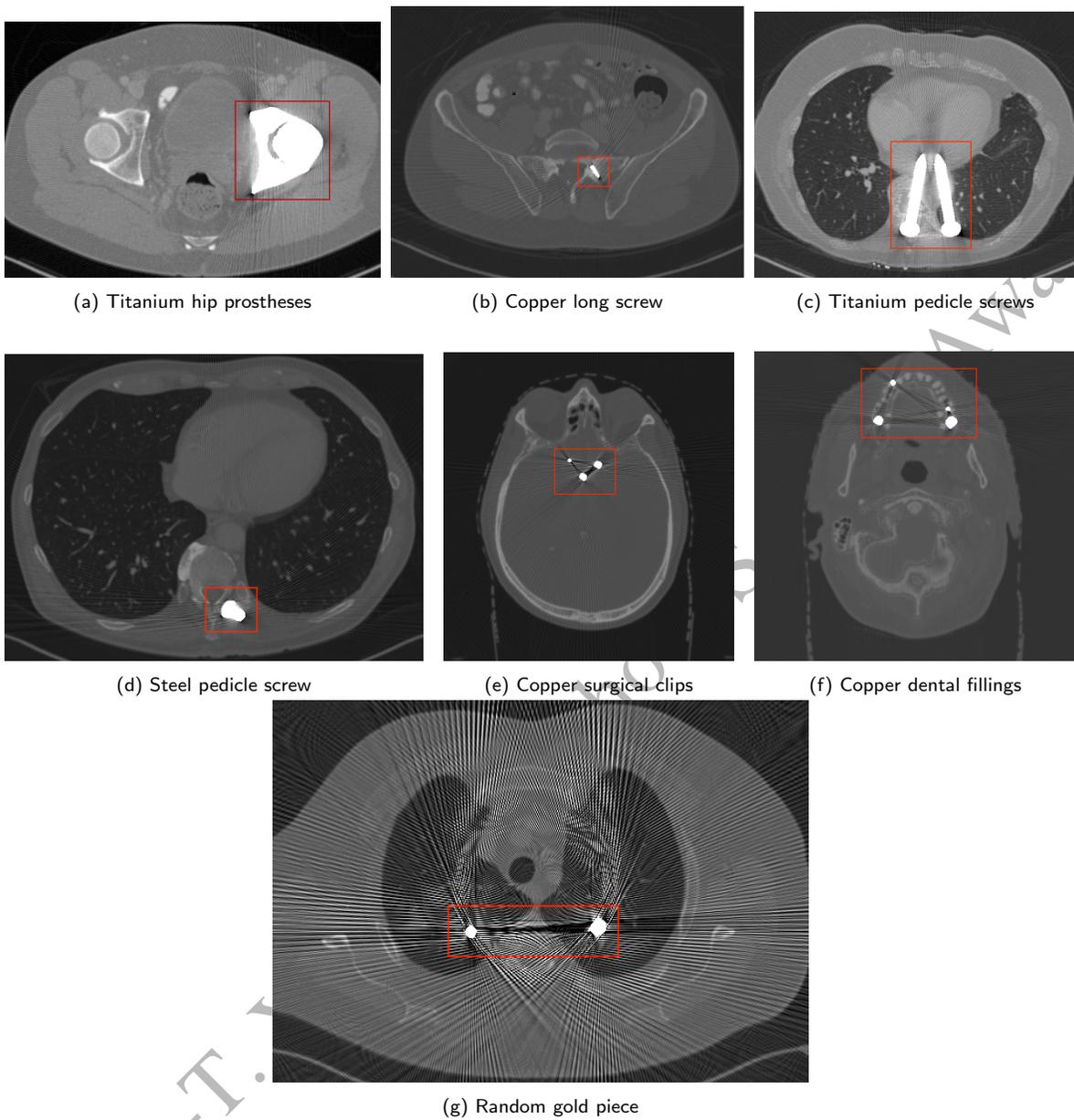
(f) Copper dental fillings

(g) Random gold piece

Figure 4: 7 reconstructed images with realistic metal implant shapes contained inside the red squares.
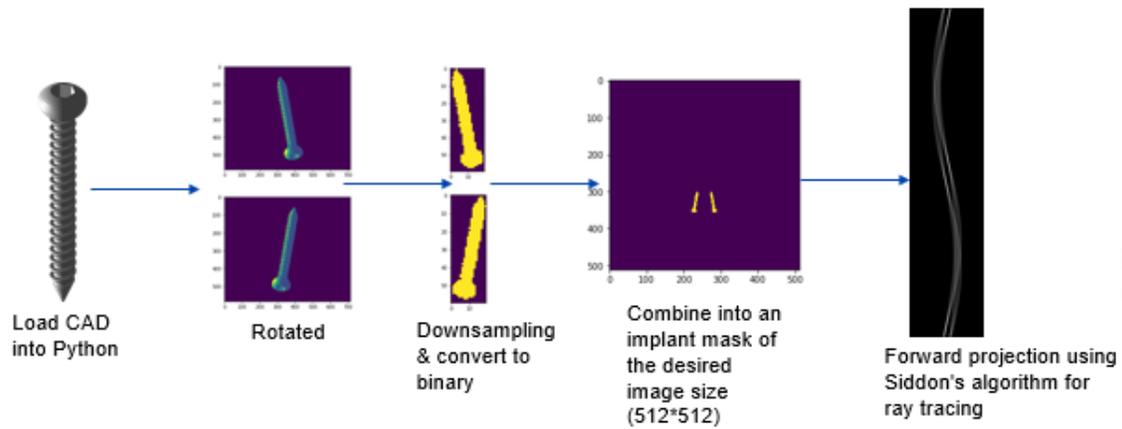
Figure 5: The process of generating a metal implant sinogram from a CAD image.

For the simulation of the second group of training data, 7 randomly generated polygonal metal inserts were assigned to a diverse range of material densities and sizes. 6 examples of these polygonal metal inserts on CT images are shown in Figure 6. To be consistent with the first group of training data and to allow a fair assessment of the effectiveness of each data generation method, the materials simulated in group 2 are correspondent with the materials simulated in group 1. The material distribution across metal implants is also correspondent with that of the group 1 training set: there are 2 titanium implants, 3 copper implants, 1 steel implant, and 1 gold implant.
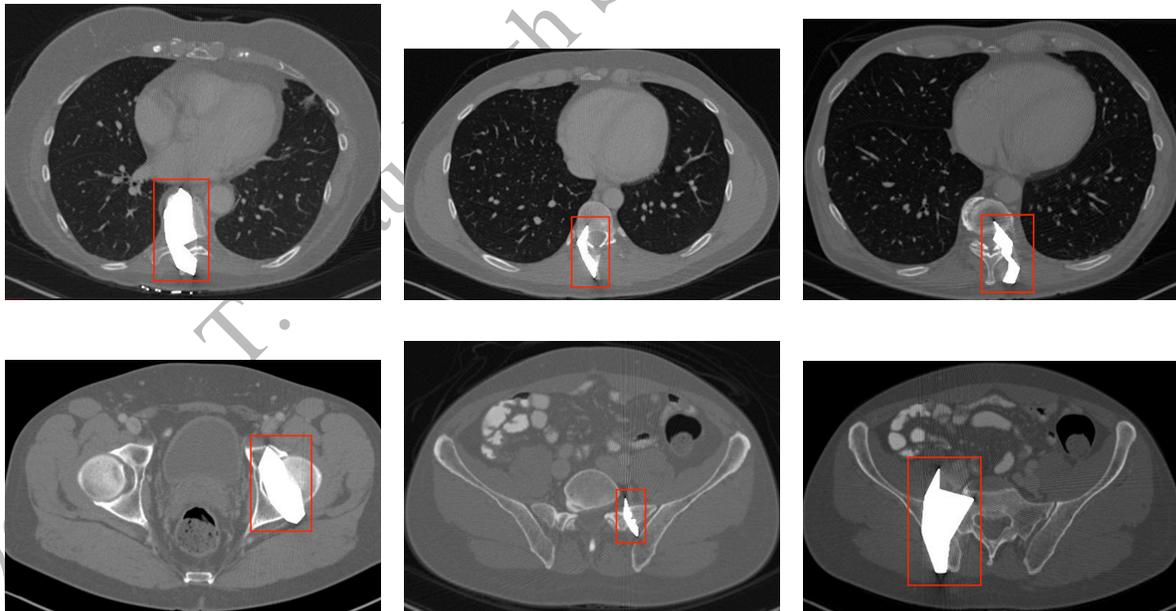

Figure 6: Examples of reconstructed images with randomly generated polygonal implants contained in the red rectangles.

The 2 models trained on both groups of training datasets were tested a small training dataset of 3 simulated metal artifact-corrupted images reconstructed from realistic metal

15

implant masks. The 3 realistic metal implant masks used for testing are: the titanium hip screw (Figure 7a), the gold pedicle screw (Figure 7c), and the steel shoulder fixation screws (Figure 7b).

(a) Titanium hip screw   (b) Steel shoulder fixation screws   (c) Gold pedicle screw
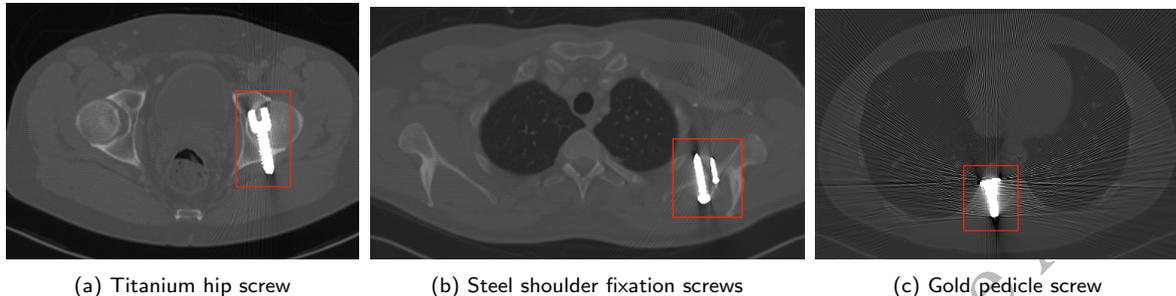
Figure 7: The testing dataset containing 3 realistic metal-artifact corrupted images. Metal implant simulations are contained in the red squares.

2.3 Proposed Network

The proposed CNN-based MAR method is inspired by the multi-material linearization method [45] that uses a polynomial function to transform the polychromatic signal to a corresponding monochromatic one. As CNNs are powerful at introducing pixel-level non-linear transformations and extracting spatial information shared by neighboring pixels (*e.g.* tissue texture, material density, *etc.*), this study replaces the polynomial function $f(x)$ in Equation 3 with a CNN to produces a corrected virtual monochromatic image (VMI) from a metal artifact-corrupted polychromatic image.

An autoencoder structure was chosen to extract the local neighborhood information from the input data. The use of symmetric contractive and expanding paths to perform encoding and decoding is inspired by the U-Net architecture [38]. The upsampling and downsampling procedures not only helps the CNN learn different levels of features from the image but also increases the computational efficiency of the model.

Inspired by the Butterfly-Net [62], The proposed method employs material decomposition pathways. To tailor the network to the MAR task, the proposed network operates on the sinogram domain to correct the data before metal artifacts appear. Because the double-entry, double-out architecture of the Butterfly-Net can only decompose bone and soft tissue, the network was modified to make for a triple-entry, single-out architecture that is capable of MAR.

The main features of the proposed convolutional autoencoder are summarized in Table 2. The advantages that these attributes provide have been verified by the previous works explored in the Related Works subsection of this paper. This is shown on Table 2, where the related works that used the techniques to achieve excellent results are listed beside the corresponding techniques.

16

Attribute	Related Works
Autoencoder architecture	[29], [39], [41], [37]
Convolutional layers	[35], [36], [30], [31], [32], [33], [34]
Material decomposition pathways	[62]
Crossover network	[62]
Use of DECT data	[23], [24], [62]

Table 2: The model's attributes and the related works that verified the advantages that these attributes.

2.3.1 Network Architecture

At the training stage of the CNN, the network takes the dual energy sinograms (2 metal artifact-corrupted sinograms taken at 80 and 140kVp tube voltages respectively) and the simulated metal path length as the training input, and takes the uncorrupted VMI as the training target.

At the application stage of the CNN, the network takes 3 inputs — the 2 dual-energy sinograms and the metal implant mask sinogram — and outputs the corrected VMI. The general architecture of the trained network is shown in Figure 8. The metal path length, if not initially known, can be acquired with the following steps: 1) reconstructing the dual-energy corrupted sinograms, 2) segmenting the metal implant from the reconstruction using the Canny edge detector [60], 3) ray tracing of the segmented metal implant to acquire its path length.
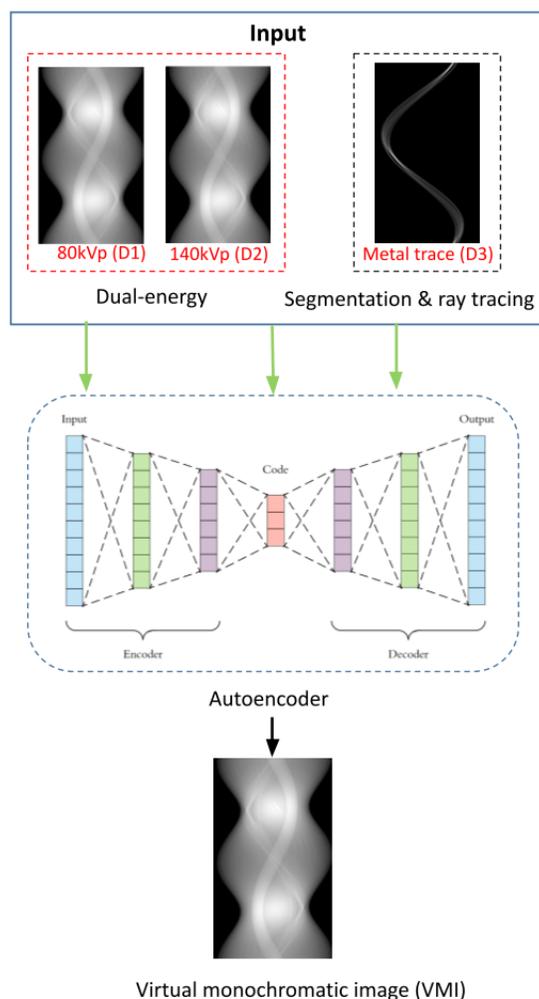
Figure 8: The triple-entry, single-out architecture of the MAR network. 3 copies of the input (corrupted dual-energy data and metal mask sinogram) are separately fed into the 3 material decomposition pathways in the convolutional autoencoder to output a VMI image.

A detailed architecture of the trained convolutional autoencoder is shown in Figure 9. The network is divided into 3 stages. Stages 1 and 2 are dedicated to material decomposition inspired by the Butterfly-Net [62]. There are 3 materials to be distinguished by the network: soft tissue, bone, and the metal of choice, which are represented by M1, M2, and M3 on Figure 9. 3 individual pathways are created to retrieve the 3 corresponding material masks in the image. Each material retrieval pathway considers 3 inputs (represented by D1, D2, and D3): 1) a metal artifact-corrupted polychromatic sinogram generated at the 80kVp tube voltage, 2) a second metal artifact-corrupted polychromatic sinogram generated at the 140kVp tube voltage, 3) the metal implant mask sinogram. The 2 sinograms generated at the clinically used DECT tube voltages of 80kVp and 140kVp are combined to reconstruct a DECT image. The path length information provided by the metal implant mask sinogram provides the CNN with spatial and contextual information necessary for auto-feature extraction.

Stage 1 involves the feature extraction that uses convolutional layers to transform the 3

18

inputs to each material pathway into a set of feature maps unique to the material. The feature maps are then fed into 3 separate streams (represented by S1, S2, and S3) and the network enters stage 2.

In stage 2, each stream uses the material feature maps learned in stage 1 to decompose the different materials in the image. The 3 separate streams are trained to have different parameters adapted to learn separate materials. Interaction between the streams, inspired by the Butterfly-Net [62], is introduced in this stage to help the convolutional layers in this stage learn about each material mask. For example, the convolutional layers of the bone pathway may take the spatial information in the metal feature map to help itself learn the spatial information about the bone mask in the image.

In stages 1 and 2, convolutional layers are used to downsample the image with a stride of 2. 3 contractive paths are used to downsample the image from the original grid size of 64×64 to a smaller grid size of 8×8. Each convolution is followed by batch normalization and a rectified linear unit (ReLU) activation function.

In stage 3, upsampling of the image is performed through 3 expanding paths that increase the grid size from 8×8 to 64×64. The 3 streams are concatenated, added, and flattened. The decomposed materials are then processed by a fully connected layer to be recombined into a single VMI output.
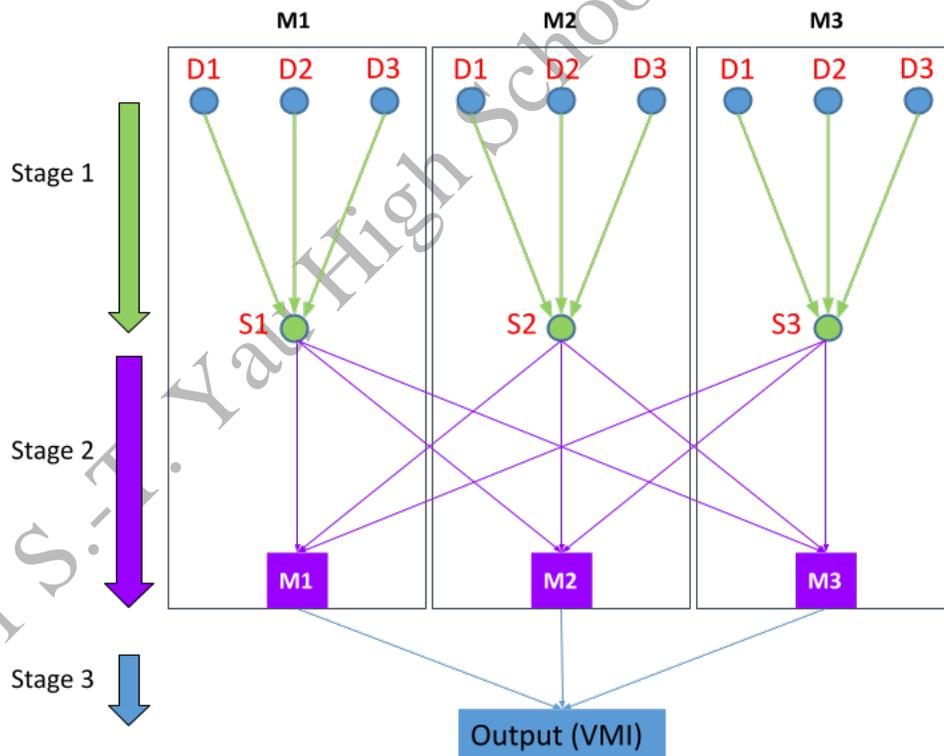

Figure 9: The structure of the autoencoder. D1, D2, and D3 represent the 3 inputs, respectively; S1, S2, S3 represent the 3 streams, respectively; M1, M2, M3 represent the 3 materials, respectively.

19

The effectiveness of the model is validated by comparing the corrected VMI with the uncorrupted ground truth VMI.

2.3.2 Training Environment

The proposed model was run on the Ubuntu operating system on the NVIDIA 2080Ti GPU hardware environment. The model was implemented with the TensorFlow and Keras frameworks in Python 3.6.

The CNN has 16 filters and uses 3×3 kernel to extract features from the input patch and map them to the target patch. The stride size is 2. The network was trained using the Adam optimizer over 200 epochs with a batch size of 32. 10% of the training set goes towards the validation set. The base learning rate was 5×10^{-4}. Batch normalization and a rectified linear unit (ReLU) activation function was introduced after every convolutional layer. Each epoch took approximately 20 minutes to complete. If there is no improvement after 10 epochs, the learning rate is reduced by a factor of 0.2. The lower bound on the learning rate is set as 1×10^{-7}.

The chosen loss function is the squared L2 norm loss, also known as mean squared error (MSE) loss. MSE is expressed as the mean of squared distances between the model's target values and predicted values. The MSE loss function seeks to minimize the value of MSE, which is calculated with the Equation below:

$$MSE = \sum_{i=1}^{N}(y_i - \hat{y_i})^2 \tag{12}$$

Where N denotes the number of samples being tested, i denotes the index of each value, $\hat{y}$ is the target value, and y is the predicted value.

2.4 Performance Metrics

The metrics chosen for model performance evaluation are Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index (SSIM).

MSE (Equation 12) represents the average squared difference between the estimated values of a pixel and the actual pixel value. The higher the MSE, the lower the quality of the reconstructed image. PSNR, measured in decibels (dB), represents the ratio between the maximum possible value (power) of a signal and the power of distorting noise that affects the quality of its representation. The equation for PSNR is presented below, where MAX_I represents the maximum possible pixel value of the image. Since the reconstructed images are 8 bits per sample, MAX_I is 255. MSE in the equation represents the MSE between the 2 images.

$$PSNR = 10 \times \log_{10}(\frac{MAX_I^2}{MSE}) \tag{13}$$

The higher the PSNR, the better the quality of the reconstructed image. Generally, a PSNR value above 35dB is considered an indicator of good image quality.

SSIM calculates the similarity between 2 images based on a weighted comparison of the luminance, contrast, and structure of the images. The maximum value of the SSIM is 1. The closer the SSIM is to 1, the better the quality of the reconstructed image.

3 Results

3.1 Model Performance

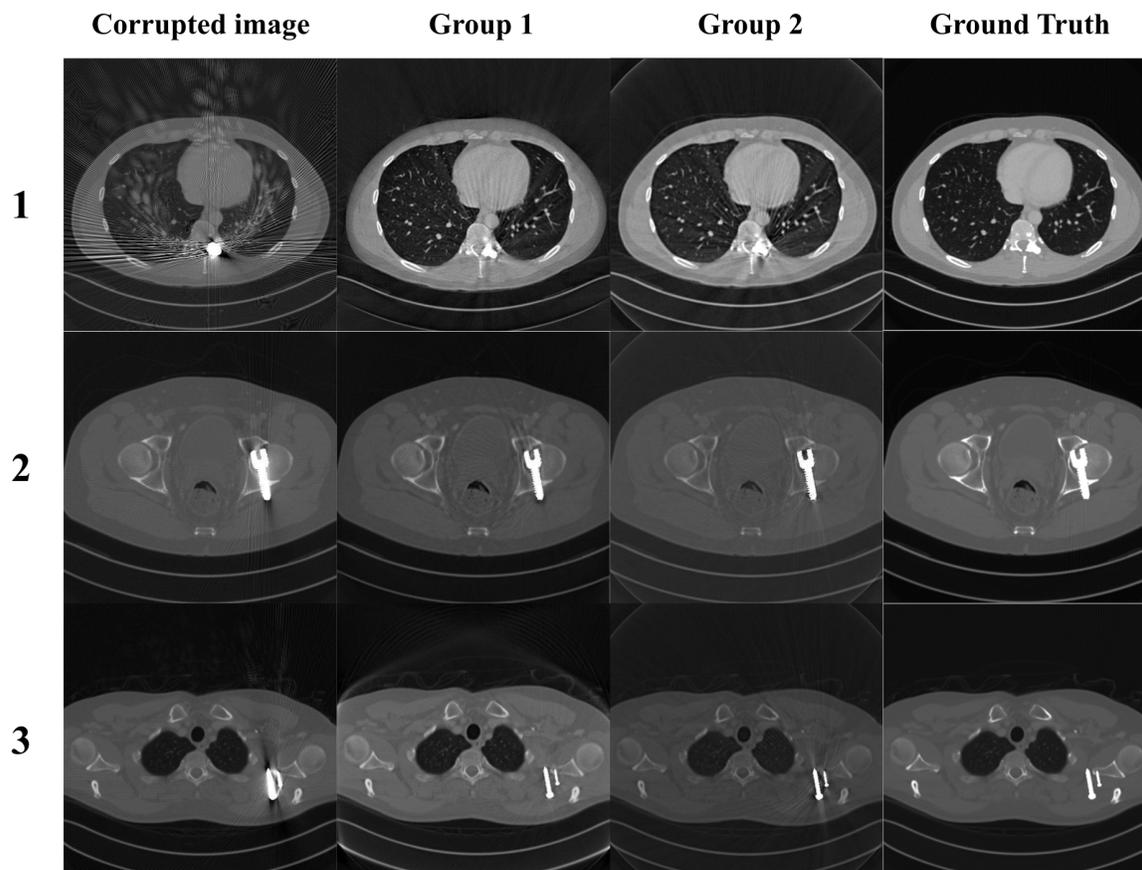| Corrupted image	Group 1	Group 2	Ground Truth

Figure 10: The respective columns of Group 1 and Group 2 present the artifact reduced images processed by the proposed CNN. The CNN of Group 1 was trained with the simulation of real metal implants, while the CNN of Group 2 was trained with randomly generated polygons. Both Groups demonstrate significant image quality improvement comparing to the corrupted image.

For ease of communication in the figures and tables in this section, each metal implant mask in the testing data is assigned to a number: 1 represents the image with a titanium hip screw mask (Figure 7a), 2 represents the image with a steel shoulder fixation screws mask (Figure 7b), 3 represents the image with a gold pedicle screw mask (Figure 7c).

The qualitative results of metal artifact correction are displayed in Figure 10. The VMI predicted by a model trained on the group 1 dataset of realistic metal implant simulations is compared to the VMI predicted by a model trained on the group 2 dataset of randomly generated polygonal metal simulations. The corrupted image (the reconstructed image generated at the 80kVp tube voltage) and the uncorrupted ground truth image are presented for reference.

Table 3 displays the quantitative evaluation of the performance of the model trained on the group 1 training set.

Metric	1	2	3	Average
MSE	9.71×10^{-6}	2.59×10^{-6}	7.46×10^{-6}	6.59×10^{-6}
PSNR (dB)	5.62×10^1	6.19×10^1	5.73×10^1	5.84×10^1
SSIM	9.89×10^{-1}	9.98×10^{-1}	9.95×10^{-1}	9.94×10^{-1}

Table 3: Quantitative performance of the model trained on the group 1 training set

Table 4 displays the quantitative evaluation of the performance of the model trained on the group 2 training set.

Metric	1	2	3	Average
MSE	5.87×10^{-6}	7.80×10^{-6}	7.27×10^{-6}	6.98×10^{-6}
PSNR (dB)	5.83×10^1	5.71×10^1	5.74×10^1	5.76×10^1
SSIM	9.94×10^{-1}	9.95×10^{-1}	9.96×10^{-1}	9.95×10^{-1}

Table 4: Quantitative performance of the model trained on the group 2 training set

As can be seen from the tables above, both models achieved outstanding performance across all evaluation metrics.

3.2 Automatic Metal Implant Generation Method

Table 5 compares the average performance of the models trained on the group 1 and group 2 training sets across all 3 testing set images. The column titled "Percentage Difference" shows the absolute value of the difference between the average score of group 1 and the average score for each evaluation metric.

Metric	Group 1 average	Group 2 average	Percentage Difference (%)
MSE	6.59×10^{-6}	6.98×10^{-6}	5.80
PSNR (dB)	5.84×10^1	5.76×10^1	1.43
SSIM	9.94×10^{-1}	9.95×10^{-1}	1.48×10^{-1}

Table 5: Quantitative performance comparison between the models trained on group 1 and group 2 training sets. The model with superior performance in each metric is highlighted in yellow.

As displayed in Figure 10 and Table 5, the performance achieved by the model trained on group 2 dataset is comparable to that achieved by the model trained on the group 1 dataset. The difference between the quantitative performance of the 2 models is relatively insignificant across all evaluation metrics. The model trained on group 2 even managed to outperform the model trained on group 1 in the SSIM metric.

4 Discussion

4.1 Summary of Contributions

This study proposed a novel end-to-end CNN algorithm for MAR in medical CT. The proposed network uses dual-energy data to yield high-quality metal artifact-reduced images with an average SSIM index higher than 0.994, an average PSNR higher than 58 dB, and an average MSE lower than 7×10^{-6}.

Compared to conventional single-energy CT, DECT's use of multiple energy spectra in data acquisition allows for a more precise measurement of energy-dependent attenuation coefficients. DECT's diagnostic advantage makes possible many clinical applications, including quantitative iodine mapping, virtual non-contrast imaging, bone differentiation, *etc*. However, even though most hospitals are equipped with DECT hardware, DECT is not widely utilized in clinical practice due to clinicians' concern that an increase in the X-ray spectra would subsequently increase the radiation dosage. This concern, however, has been proven unfounded: DECT does not increase the radiation dosage compared to single-energy CT [4]. With the proposed convolutional autoencoder, this study has demonstrated DECT's potential in MAR.

Along with the MAR network, an innovative synthetic data generation method was proposed to automate the metal implant simulation process by replacing manually segmented realistic metal implant shapes with randomly generated polygons. Not only does the model trained on data simulated using polygons achieve comparable performance to the model trained on the same number of realistic metal implant simulations, but the improved efficiency of the metal implant generation process can boost the generalizability of the CNN by allowing for easy collection of large datasets.

4.2 Limitations and Future Work

The major source of weakness of the proposed model is its small training dataset size. A small training dataset makes it difficult to optimize performance, especially for deep learning algorithms as their performance is heavily reliant on dataset size. However, one of the goals of this study was to create a model that minimizes the need for a large training dataset. The proposed network certainly achieved this goal. The excellent results achieved by the model trained on a dataset of merely 390 instances reflect the tremendous potential of the proposed network.

Another potential drawback of the CNN is posed by its requirement for knowledge of the metal implant's path length. To acquire the path length, the metal implant mask needs to be segmented. If photon starvation (insufficient photons reaching the detector) occurs, the metal implant would be incorrectly segmented. In the future, a method needs to be developed to correct the signal corrupted by photon starvation.

Lastly, this study was limited by the absence of comparative performance analysis. Although the original intention of this study was to compare the proposed MAR model with existing state-of-the-art MAR models, there were no available open source codes of MAR algorithms to implement and compare performance with. Future work on this topic would need to include comparison with other methods.

References

[1] Barrett, Julia F., and Nicholas Keat. "Artifacts in CT: Recognition and Avoidance." *RadioGraphics*, vol. 24, no. 6, Nov. 2004, pp. 1679–1691, pubs.rsna.org/doi/10.1148/rg.246045065, 10.1148/rg.246045065.

[2] Sprawls, Perry. *Physical Principles of Medical Imaging*. Madison, Wis., Medical Physics Pub, 1995.

[3] Kak, Avinash C, and Malcolm Slaney. *Principles of computerized tomographic imaging*. Society for Industrial and Applied Mathematics, 2001.

[4] Henzler, Thomas, et al. "Dual-energy CT: radiation dose aspects." *American Journal of Roentgenology* 199.5_supplement (2012): S16-S25.

[5] Giantsoudi, Drosoula et al. "Metal artifacts in computed tomography for radiation therapy planning: dosimetric effects and impact of metal artifact reduction." *Physics in medicine and biology*; vol. 62,8 (2017): R49-R80. doi:10.1088/1361-6560/aa5293

[6] CAO, W., et al. "Comparative Performance Assessment of Beam Hardening Correction Algorithms Applied on Simulated Data Sets." *Journal of Microscopy*, vol. 272, no. 3, 8 Aug. 2018, pp. 229–241, 10.1111/jmi.12746. Accessed 4 Aug. 2021.

[7] W. A. Kalender, R. Hebel, and J. Ebersberger, "Reduction of CT artifacts caused by metallic implants," *Radiology*, vol. 164, pp. 576–577, Aug. 1987.

[8] E. Meyer, R. Raupach, M. Lell, B. Schmidt, and M. Kachelrieß, "Normalized metal artifact reduction (NMAR) in computed tomography," *Med. Phys.*, vol. 37, no. 10, pp. 5482–5493, Oct. 2010.

[9] McCollough, Cynthia H., et al. "Dual- and Multi-Energy CT: Principles, Technical Approaches, and Clinical Applications." *Radiology*, vol. 276, no. 3, Sept. 2015, pp. 637–653, 10.1148/radiol.2015142631.

[10] J. M. Verburg, and J. Seco, "CT metal artifact reduction method correcting for beam hardening and missing projections," *Physics in Medicine Biology*, vol. 57, no. 9, pp. 2803-2818, 2012.

[11] M. Bal and L. Spies, "Metal artifact reduction in CT using tissue-class modeling and adaptive prefiltering," *Med. Phys.*, vol. 33, no. 8, pp. 2852–2859, 2006.

[12] D. Prell, Y. Kyriakou, M. Beister, and W. A. Kalender, "A novel forward projection-based metal artifact reduction method for flat-detector computed tomography," *Phys. Med. Biol.*, vol. 54, no. 21, pp. 6575–6591, Oct. 2009.

[13] J. Wang, S. Wang, Y. Chen, J. Wu, J.-L. Coatrieux, and L. Luo, "Metal artifact reduction in CT using fusion based prior image," *Med. Phys.*, vol. 40, no. 8, p. 081903, 2013.

[14] Y. Zhang and X. Mou. (2014). "Metal artifact reduction based on the combined prior image." [Online]. Available: https://arxiv. org/abs/1408.5198

[15] L. Gjesteby et al., "Metal artifact reduction in CT: Where are we after four decades?" *IEEE Access*, vol. 4, pp. 5826–5849, 2016.

[16] Je. Y. Huang et al., "An evaluation of three commercially available metal artifact reduction methods for CT imaging," *Phys. Med. Biol.*, vol. 60, no. 3, pp. 1047–1067, 2015.

[17] A. Mouton, N. Megherbi, K. Van Slambrouck, J. Nuyts, and T. P. Breckon, "An experimental survey of metal artefact reduction in computed tomography," *J. X-Ray Sci. Technol.*, vol. 21, no. 2, pp. 193–226, 2013.

[18] G. Wang, D. L. Snyder, J. A. O'Sullivan, and M. W. Vannier, "Iterative deblurring for CT metal artifact reduction," *IEEE Trans. Med. Imag.*, vol. 15, no. 5, pp. 657–664, Oct. 1996.

[19] G. Wang, M. W. Vannier, and P.-C. Cheng, "Iterative X-ray cone-beam tomography for metal artifact reduction and local region reconstruction," *Microscopy Microanal.*, vol. 5, no. 1, pp. 58–65, 1999.

[20] X. Zhang, J. Wang, and L. Xing, "Metal artifact reduction in X-ray computed tomography (CT) by constrained optimization," *Med. Phys.*, vol. 38, no. 2, pp. 701–711, 2011.

[21] Y. Zhang, X. Mou, and H. Yan, "Weighted total variation con- strained reconstruction for reduction of metal artifact in CT," in Proc. *IEEE Nucl. Sci. Symp. Med. Imag. Conf.* (NSS/MIC), Oct./Nov. 2010, pp. 2630–2634.

[22] C. Lemmens, D. Faul, and J. Nuyts, "Suppression of metal artifacts in CT using a reconstruction procedure that combines MAP and projection completion," *IEEE Trans. Med. Imag.*, vol. 28, no. 2, pp. 250–260, Feb. 2009.

[23] Lehmann, L. A., et al. "Generalized Image Combinations in Dual kVp Digital Radiography." *Medical Physics*, vol. 8, no. 5, Sept. 1981, pp. 659–667, 10.1118/1.595025. Accessed 5 Aug. 2021.

[24] Alvarez, R E, and A Macovski. "Energy-Selective Reconstructions in X-Ray Computerised Tomography." *Physics in Medicine and Biology*, vol. 21, no. 5, 1 Sept. 1976, pp. 733–744, 10.1088/0031-9155/21/5/002. Accessed 5 Aug. 2021.

[25] Vetter, J. R., et al. "Evaluation of a Prototype Dual-Energy Computed Tomographic Apparatus. II. Determination of Vertebral Bone Mineral Content." *Medical Physics*, vol. 13, no. 3, May 1986, pp. 340–343, 10.1118/1.595951. Accessed 5 Aug. 2021.

[26] Y. Zhang and X. Mou, "Metal artifact reduction based on beam hardening correction and statistical iterative reconstruction for X-ray computed tomography," *Proc. SPIE*, vol. 8668, p. 86682O, Mar. 2013.

[27] Zhang, Yanbo, et al. "A Hybrid Metal Artifact Reduction Algorithm for X-Ray CT." *Medical Physics*, vol. 40, no. 4, 21 Mar. 2013, p. 041910, 10.1118/1.4794474. Accessed 5 Aug. 2021.

[28] Gjesteby, Lars, et al. "Deep Learning Methods to Guide CT Image Reconstruction and Reduce Metal Artifacts." *Medical Imaging* 2017: Physics of Medical Imaging, 9 Mar. 2017, 10.1117/12.2254091. Accessed 5 Aug. 2021.

[29] Park, Hyoung Suk, et al. "CT Sinogram-Consistency Learning for Metal-Induced Beam Hardening Correction." *Medical Physics*, vol. 45, no. 12, 8 Nov. 2018, pp. 5376–5384, 10.1002/mp.13199. Accessed 5 Aug. 2021.

[30] L. Gjesteby et al., "Reducing metal streak artifacts in CT images via deep learning: Pilot results," in *Proc. 14th Int. Meeting Fully Three-Dimensional Image Reconstruction Radiol. Nucl. Med.*, 2017, pp. 611–614.

[31] W. Du, H. Chen, Z. Wu, H. Sun, P. Liao, and Y. Zhang, "Stacked competitive networks for noise reduction in low-dose CT," *PLoS ONE*, vol. 12, no. 12, p. e0190069, 2017.

[32] Chen, Hu, et al. "LEARN: Learned Experts' Assessment-Based Reconstruction Network for Sparse-Data CT." *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, June 2018, pp. 1333–1347, 10.1109/tmi.2018.2805692. Accessed 5 Aug. 2021.

[33] Y. Liu and Y. Zhang, "Low-dose CT restoration via stacked sparse denoising autoencoders," *Neurocomputing*, vol. 284, pp. 80–89, Apr. 2018.

[34] H. Chen et al., "Low-dose CT via convolutional neural network," *Biomed. Opt. Exp.*, vol. 8, no. 2, pp. 679–694, 2017.

[35] Zhang, Yanbo, and Hengyong Yu. "Convolutional Neural Network Based Metal Artifact Reduction in X-Ray Computed Tomography." *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, 1 June 2018, pp. 1370–1381, pubmed.ncbi.nlm.nih.gov/29870366/, 10.1109/TMI.2018.2823083. Accessed 5 Aug. 2020.

[36] Huang, Xia, et al. "Metal Artifact Reduction on Cervical CT Images by Deep Residual Learning." *BioMedical Engineering OnLine*, vol. 17, no. 1, 27 Nov. 2018, 10.1186/s12938-018-0609-y. Accessed 5 Aug. 2021.

[37] Liao, Haofu, et al. "ADN: Artifact Disentanglement Network for Unsupervised Metal Artifact Reduction." *IEEE Transactions on Medical Imaging*, vol. 39, no. 3, Mar. 2020, pp. 634–643, 10.1109/tmi.2019.2933425. Accessed 5 Aug. 2021.

[38] Ronneberger, Olaf, et al. "U-Net: Convolutional Networks for Biomedical Image Segmentation." ArXiv:1505.04597 [Cs], 18 May 2015, arxiv.org/abs/1505.04597v1. Accessed 5 Aug. 2021.

[39] M. Du, K. Liang and Y. Xing, "Reduction of metal artefacts in CT with Cycle-GAN," *2018 IEEE Nuclear Science Symposium and Medical Imaging Conference Proceedings (NSS/MIC)*, 2018, pp. 1-3, doi: 10.1109/NSSMIC.2018.8824544.

[40] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A. Efros. "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks", in *IEEE International Conference on Computer Vision (ICCV)*, 2017.

[41] Lyu, Yuanyuan, et al. "U-DuDoNet: Unpaired Dual-Domain Network for CT Metal Artifact Reduction." ArXiv:2103.04552 [Cs, Eess], 8 Mar. 2021, arxiv.org/abs/2103.04552. Accessed 5 Aug. 2021.

[42] Jiang Hsieh. *Computed Tomography : Principles, Design, Artifacts, and Recent Advances.* Hoboken, Nj, Wiley Interscience, 2009.

[43] Zhou, W., et al. "Comparison of a Photon-Counting-Detector CT with an Energy-Integrating-Detector CT for Temporal Bone Imaging: A Cadaveric Study." *American Journal of Neuroradiology*, vol. 39, no. 9, 9 Aug. 2018, pp. 1733–1738, 10.3174/ajnr.a5768. Accessed 7 Aug. 2021.

[44] Herman, Gabor T., and Sushma S. Trivedi. "A Comparative Study of Two Postreconstruction Beam Hardening Correction Methods." *IEEE Transactions on Medical Imaging*, vol. 2, no. 3, Sept. 1983, pp. 128–135, 10.1109/tmi.1983.4307626. Accessed 7 Aug. 2021.

[45] Joseph, Peter M., and Robin D. Spital. "A Method for Correcting Bone Induced Artifacts in Computed Tomography Scanners." *Journal of Computer Assisted Tomography*, vol. 2, no. 1, Jan. 1978, pp. 100–108, 10.1097/00004728-197801000-00017. Accessed 2 Apr. 2020.

[46] Hubbell, J, and S Seltzer. "X-Ray Mass Attenuation Coefficients." , 2004. *NIST Standard Reference Database 126.*

[47] Yan, Ke, et al. "DeepLesion: Automated Mining of Large-Scale Lesion Annotations and Universal Lesion Detection with Deep Learning." *Journal of Medical Imaging*, vol. 5, no. 03, 20 July 2018, p. 1, 10.1117/1.jmi.5.3.036501. Accessed 16 Nov. 2019.

[48] Punnoose, J., et al. "Technical Note: Spektr 3.0-A Computational Tool for X-Ray Spectrum Modeling and Analysis." *Medical Physics*, vol. 43, no. 8Part1, 21 July 2016, pp. 4711–4717, www.osti.gov/biblio/22689405-technical-note-spektr-computational-tool-ray-spectrum-modeling-analysis, 10.1118/1.4955438. Accessed 11 Aug. 2021.

[49] W. van Aarle, W. J. Palenstijn, J. Cant, E. Janssens, F. Bleichrodt, A. Dabravolski, J. De Beenhouwer, K. J. Batenburg, and J. Sijbers, "Fast and Flexible X-ray Tomography Using the ASTRA Toolbox", *Optics Express*, 24(22), 25129-25147, (2016), http://dx.doi.org/10.1364/OE.24.025129

[50] Parakh, Anushri, et al. "Dual-Energy CT Images: Pearls and Pitfalls." *RadioGraphics*, vol. 41, no. 1, Jan. 2021, pp. 98–119, 10.1148/rg.2021200102. Accessed 27 Aug. 2021.

[51] Mahesh, Mahadevappa. "[I156] Why Is Dual Energy CT (DECT) Still Not Used Widely in Clinics?" *Physica Medica*, vol. 52, Aug. 2018, p. 60, 10.1016/j.ejmp.2018.06.228. Accessed 14 Apr. 2019.

[52] Browne, Matthew, and Saeed Shiry Ghidary. "Convolutional Neural Networks for Image Processing: An Application in Robot Vision." *Lecture Notes in Computer Science*, 2003, pp. 641–652, 10.1007/978-3-540-24581-0_55. Accessed 10 Sept. 2020.

[53] Mohsen, Heba, et al. "Classification Using Deep Learning Neural Networks for Brain Tumors." *Future Computing and Informatics Journal*, vol. 3, no. 1, June 2018, pp. 68–71,

www.sciencedirect.com/science/article/pii/S2314728817300636, 10.1016/j.fcij.2017.12.001. Accessed 29 Aug. 2021.

[54] Lossau (née Elss), T., et al. "Learning Metal Artifact Reduction in Cardiac CT Images with Moving Pacemakers." *Medical Image Analysis*, vol. 61, Apr. 2020, p. 101655, www.sciencedirect.com/science/article/abs/pii/S1361841520300220, 10.1016/j.media.2020.101655. Accessed 30 Aug. 2021.

[55] Gardner, Chester S., and George Papen. "Optical Communications." *Reference Data for Engineers*, 2002, pp. 22–122–28, www.sciencedirect.com/topics/physics-and-astronomy/attenuation-coefficients, 10.1016/b978-075067291-7/50024-8. Accessed 30 Aug. 2021.

[56] H. S. Park, D. Hwang, and J. K. Seo, "Metal artifact reduction for polychromatic x-ray CT based on a beam-hardening corrector," *IEEE Transactions on Medical Imaging*, vol. 35, no. 2, pp. 480–487, 2016.

[57] J. Hsieh, R. C. Molthen, C. A. Dawson, and R. H. Johnson, "An iterative approach to the beam hardening correction in cone beam CT," *Medical Physics*, vol. 27, no. 1, pp. 23–29, 2000.

[58] Y. Zhang, X. Mou, and S. Tang, "Beam hardening correction for fan- beam CT imaging with multiple materials," in 2010 *IEEE Nuclear Science Symposium and Medical Imaging Conference* (2010 NSS/MIC), 2010, pp. 3566–3570.

[59] M. Kachelrieß, O. Watzke, and W. A. Kalender, "Generalized multi- dimensional adaptive filtering for conventional and spiral single-slice, multi-slice, and cone-beam CT," *Medical physics*, vol. 28, no. 4, pp. 475–490, 2001.

[60] Canny, J., A Computational Approach To Edge Detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698, 1986.

[61] Zeng, Gengsheng Lawrence. Medical Image Reconstruction. Berlin, Heidelberg, *Springer Berlin Heidelberg*, 2010. Accessed 29 Aug. 2021.

[62] Zhang, Wenkun, et al. "Image Domain Dual Material Decomposition for Dual-EnergyCTusing Butterfly Network." Medical Physics, vol. 46, no. 5, Apr. 2019, pp. 2037–2051, pubmed.ncbi.nlm.nih.gov/30883808/, 10.1002/mp.13489. Accessed 29 Aug. 2021.