Journal of Innovations in Computer Science and Engineering
Shahid Beheshti Unversity

June 2025, Volume 3, Issue 1

Automatic Multi-Class Cardiovascular Magnetic Resonance Image Quality Assessment using Unsupervised Domain Adaptation in Spatial and Frequency Domains

Shahabedin Nabavi✉, Code Orcide: 0000-0001-7240-0239
Faculty of Computer Science and Engineering, Shahid Beheshti University, Tehran, Iran. s_nabavi@sbu.ac.ir
Hossein Simchi, **Code Orcide:** 0009-0002-1392-1426
Faculty of Computer Science and Engineering, Shahid Beheshti University, Tehran, Iran. hsimchi74@gmail.com
Mohsen Ebrahimi Moghaddam, **Code Orcide:** 0000-0002-7391-508X
Faculty of Computer Science and Engineering, Shahid Beheshti University, Tehran, Iran. m_moghadam@sbu.ac.ir
Ahmad Ali Abin, **Code Orcide:** 0000-0002-0916-0348
Faculty of Computer Science and Engineering, Shahid Beheshti University, Tehran, Iran. a_abin@sbu.ac.ir
Alejandro F. Frangi, **Code Orcide:** 0000-0002-2675-528X
Division of Informatics, Imaging and Data Sciences, Schools of Computer Science and Health Sciences, The University of Manchester, Manchester, UK.
Medical Imaging Research Center (MIRC), Electrical Engineering and Cardiovascular Sciences Departments, KU Leuven, Leuven, Belgium.
Alan Turing Institute, London, UK. alejandro.frangi@manchester.ac.uk

Abstract

Population imaging studies rely on good-quality medical imagery before quantifying downstream images. This study provides an automated approach for image quality assessment (IQA) from cardiovascular magnetic resonance (CMR) imaging at scale. We identify four common CMR imaging artefacts: respiratory motion, cardiac motion, Gibbs ringing, and aliasing. Four datasets, including UK Biobank, York University (YU), the Universidad Carlos III (UCIII) and CMR-Tehran, were used to perform the experiments. This study proposes two deep-learning models for CMR IQA in spatial and frequency domains. The presented spatial-domain model also has domain adaptation. The accuracies of supervised 4-fold cross-validation experiments for UK Biobank, YU, UCIII and CMR-Tehran datasets are 99.41%, 75.78%, 89.46% and 67.87% for the spatial-domain and 87.46%, 63.76%, 80.25% and 58.48% for the frequency-domain. Domain adaptation results, considering UK Biobank as the source set and YU, UCIII and CMR-Tehran as the target sets, show the domain shift gap coverage between the datasets to the extent of +11.91%, +3.93% and +16.57%, respectively. Besides, by training and testing the spatial-domain model on 30,125 images from the UK Biobank, an accuracy of 89.56% was obtained in a training time of 394.80 seconds. Meanwhile, the frequency-domain model with training and testing on 180,750 images achieves an accuracy of 87.99% in a training time of 255.04 seconds. Thus, the frequency-domain model can achieve almost the same accuracy yet 1.548 times faster than the spatial model. The proposed models can detect four common CMR imaging artefacts by receiving images or the corresponding k-spaces.

KEYWORDS: Artefact, Cardiovascular magnetic resonance imaging, Deep learning, Domain Adaptation, Image quality assessment.

1- Introduction

Cardiovascular magnetic resonance (CMR) imaging has many clinical applications as a powerful non-invasive diagnostic tool. This imaging modality can assess cardiac function and support the diagnosis of cardiovascular diseases, including coronary artery disease, cardiomyopathy, and congenital and valvular disease [1]. CMR imaging is the gold standard in many diagnostic and therapeutic applications [2, 3]. In addition, CMR is used in most cardiovascular population imaging studies, given its optimal trade-off of non-invasiveness and accuracy. Imaging artefacts hinder routine quantitative analysis of CMR, amongst other factors. These artefacts affect the cardiac image analysis, rendering cardiac anatomy or functional parameters inaccurate for diagnostic or epidemiological studies. The increasingly wider availability of CMR expertise in mainstream cardiology and the

Submit Date: 2025-03-09
Revise Date: 2025-07-21
Accept Date: 2025-09-17
✉ Corresponding author

development of novel MRI protocols have reduced the incidence of these artefacts. Yet, their presence remains unavoidable, and visual determination of image quality is impractical in busy cardiovascular imaging departments. Hence, automated approaches to detect and identify these artefacts remain highly desirable.

After each image acquisition and over an extended period, image quality needs to be assessed to ensure optimal diagnostic value and minimal image quality drifts over time. Human quality assessment of CMR images can be time-consuming and costly. The many images acquired during an imaging examination require automatic machine learning methods for image quality assessment. Chow and Paramesran reviewed the methods of assessing the quality of medical images and specifically predicted that the next generation of methods would be no-reference [4]. Availability of the reference image means that the quality of each received image is compared to the reference image. The main challenge in the quality assessment of medical images is the impossibility of accessing this reference image, so estimating the quality of medical images is generally established without reference.

In a study by Tarroni et al. [5], an automated process for determining the quality of short and long-axis CMR images is proposed, examining image quality related to full heart coverage in the images, motion artefact, and contrast estimation. That study has been performed on two datasets, one with 3000 samples for training and testing processes and 100 samples for testing based on random forests. That study shows the sensitivity and specificity of 88 and 99% for the diagnosis of full cardiac coverage and 85 and 99% for the diagnosis of motion artefact. In another study [6], the same method as [5] has been evaluated on a large-scale subset of the UK Biobank. The results on 19,265 short-axis cine stacks from the UK Biobank show that up to 14.2% have suboptimal coverage and up to 16% have inter-slice motion. In the study of Oksuz et al. [7], motion artefacts due to breathing and cardiac movements were investigated. In that study, k-space manipulation has been used to increase the volume of distorted data, and a CNN architecture has been used to learn the automatic detection model. The results on 3510 CMR images show an area under the receiver operating characteristic (ROC) curve (AUC) of 0.89. In a study by Zhang et al. [8], the problem of full left ventricular coverage using CNNs has been investigated in a data set of over 5,000 cases, with an error rate of less than 5%. In that study, short-axis cine CMR images have been received as input. Two parallel CNN architectures have been used to determine the presence or absence of images related to the apex and the basal of the heart. The main innovation of that study is using a particular layer in the end part of the proposed architecture to improve the classification of images, which has been introduced as a Fisher discriminative layer. This layer tries to minimise within-class variance and maximise the distance of between-class means. In another study [9], the use of an adversarial learning approach based on CNNs to detect the entire left and right ventricular coverage of the heart has been investigated. That method has been studied on three datasets, and the results show the superiority of the method over previous approaches. In another study [10], generative adversarial networks have been evaluated to detect full left ventricular coverage in a dataset consisting of over 6,000 samples. The accuracy obtained in that study has been reported to be about 90%. In a study by Osadebey et al. [11], image quality assessment in the brain and cardiovascular images has been investigated. In that study, 16 volumes of CMR images have been used, and the method is based on the handmade feature extraction of four types of additive noise. Several other studies have also addressed CMR image quality assessment, including [12, 13]. A summary of these related studies is given in Table S1 (See supplementary material).

This study proposes an automated CMR image quality assessment method to detect different artefacts. Besides image quality control, it also reveals the type of artefact. Thus, if the imaging needs to be repeated, it makes the imaging technicians aware of the source of the problem.

The contributions of this study are:

(1) An automated image quality assessment approach is proposed that handles four common types of CMR artefacts. Respiratory motion, cardiac motion, Gibbs ringing, and aliasing artefacts are examined. Considering the four common artefacts in CMR images, this study outperforms previous studies regarding the number of examined artefacts.
(2) In this approach, CMR images with different views, including short-axis cine CMR, two, three, and four-chamber long-axis images, are considered input. We seek to obtain more generality for the proposed models from this point of view.
(3) The proposed method can be used in both spatial and frequency domains. Therefore, the frequency domain model can assess the quality based on receiving k-space. The k-spaces in this study are obtained by Fourier transformation of the corresponding magnitude images and are not the scanner-acquired k-spaces.
(4) Increasing the speed of training and testing using the proposed frequency domain model is another contribution of this study.
(5) The approach is evaluated on several datasets to avoid biasing the assessment and understand the generalisability of the process. Since access to data labels is usually limited, a method is proposed to generate a training set by modelling imaging artefacts using k-space manipulation, producing corrupted images. A deep domain adaptation method in the spatial domain is then presented and evaluated for unavailable data labels in some datasets. These proposed methods are assessed in supervised, learning from limited training data, and unsupervised manners.

2- Materials and Methods

In this section, the datasets used in the study are first described. Then the methods of adding artefacts to CMR images by manipulating the k-space to generate four types of artefacts, including respiratory and cardiac motion, aliasing and Gibbs ringing, are explained. This section also presents deep learning-based models for identifying CMR image artefacts. These models are proposed in both spatial and frequency domains, and details of all models are described. An unsupervised domain adaptation method is presented for conditions where the data labels are not accessible. In the end, the experimental setups are described to explain the defined experiments and evaluation metrics.

2-1- Dataset description

The main database of the current study is the CMR image database from the UK Biobank. Over 35,000 2D slices from over 6,000 subjects of this database were used to generate a dataset including CMR images corrupted by synthetic artefacts. CMR images were acquired using a wide clinical bore 1.5*T* MR system (MAGNETOM Aera, Syngo Platform VD13A, Siemens

Healthcare, Erlangen, Germany) with an 18 channel anterior body surface coil (45 mT/m and 200 $T/m/s$ gradient system). 2D cine balanced steady-state free precession (bSSFP) short-axis and long-axis images have in-plane spatial resolution $1.8 \times 1.8\ mm$, image size 198×208, slice thickness 8 mm and slice gap 2 mm. More information about the imaging protocol is available in [14].

Besides UK Biobank, data from three other datasets have been used. One of these datasets is the cardiac MRI dataset from York University (YU). A total of 7980 short and long-axis images have been acquired from 33 subjects. Imaging has been performed with a GE Genesis Signa MR scanner using the fast imaging employing steady-state acquisition (FIESTA) scanning protocol. The subjects studied in this research are all under 18 years old. The size of each slice is 256×256 pixels. In the short-axis view, the number of frames is 20; in the long-axis view, the number of slices varies between 8 and 15. The slice gap is also between 6 and 13 mm. Details related to this dataset can be found in [15]. The Universidad Carlos III (UCIII) dataset [16], containing 32 short-axis cardiac cine MR images of four small animals, is another dataset used in this study. A $7T$ Bruker Biospec $70/20$ scanner has been used to scan these self-gated rat cardiac cine sequences (IntraGateFLASH). The number of acquired frames is 8 with an image size of 192×192. slice thickness is 1.2 mm. Finally, another dataset including short and long-axis CMR images was collected from Tehran hospitals. 2620 images of 21 unknown subjects have been gathered in this dataset. The imaging parameters are different between CMR image volumes in this dataset due to its multicentre nature. Patient information and images related to this database will remain confidential. We intentionally selected datasets with varying sizes, qualities, and subject types to introduce a significant domain shift, allowing us to assess the model's robustness and generalization under challenging, real-world conditions.

For data augmentation in YU, UCIII, and CMR-Tehran datasets with fewer data, horizontal and vertical flips, random rotation with padding and brightness-changing methods have been used. The angle of rotation and the number of brightness changes were considered random to diversify the data further. Data augmentation methods are used to increase the size of the training sets.

2-2- Artefacted CMR image ground truth via k-space manipulation

Due to hardware and software restrictions, CMR images may be associated with artefacts. These artefacts can seriously affect image quality and thus the ability to make a proper diagnosis. According to the published statistics, motion and Gibbs artefacts are the most common in magnetic resonance imaging [17]. Specifically, respiratory and cardiac motion, Gibbs ringing, and aliasing artefacts are the most common in CMR imaging [18]. For this reason, we investigated these types of artefacts in this study.

Subjectively determining the artefact in the CMR images is a time consuming and laborious task that requires several human observers. Since this study needs corrupted images, we used the k-space degradation methods to generate synthetic images. Therefore, long-axis and short-axis cine CMR images acquired using Cartesian sampling were used to add respiratory and cardiac motion, aliasing and Gibbs ringing artefacts. These sections explain how to corrupt images using k-space manipulation.

2-2-1- Respiratory motion artefacts

We use a Cartesian manipulation approach to make corrupted CMR images with the respiratory artefact through k-space manipulation, considering that imaging datasets used in this study are acquired using Cartesian sampling. According to the method presented in Lorch et al. [19] study, we leverage a translation with a sinusoidal pattern on the reference images. First, the reference image is subjected to a 1D translation along the height or width. Then, due to Cartesian sampling, the fast Fourier transform is used to transfer the reference and the translated image into k-space. In the next step, some lines from the k-space of the reference image are replaced with corresponding lines related to the translated image. Finally, the modified k-space is reconstructed into the image space using the inverse fast Fourier transform, and we see the appearance of artefacts in the reconstructed image.

Translation in the height or width of the reference image to generate images with different severities of the artefact is considered from 1 to 5 pixels as a random selection. Also, 10 to 25 lines from the k-space are randomly selected for the replacement to make distorted images with different severities of the artefact. This method of generating degraded images is also used in the study by Oksuz et al. [7]. Figure 1 shows how to develop a corrupted slice with this method.

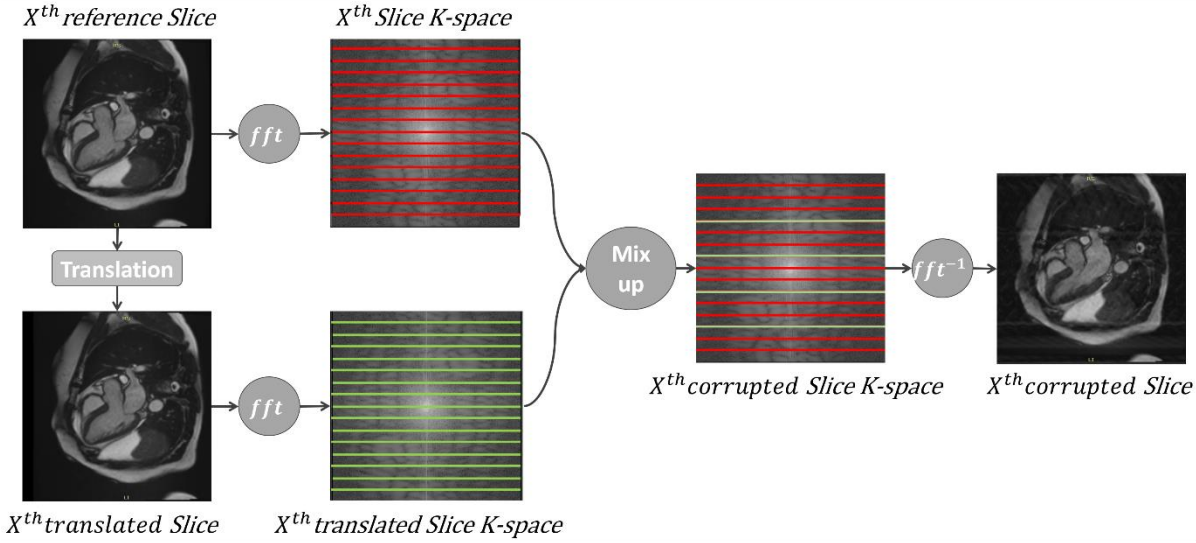


Figure 1. k-space manipulation to generate synthetic respiratory artefacts.

2-2-2- Cardiac motion artefacts

Similar to [7], short-axis cine CMR images need to be used to generate degraded images with cardiac motion artefacts. For this purpose, a temporal sequence is separated from the short-axis cine images that fully cover the heart from the basal to the apex. The images in this sequence are then transferred to the corresponding k-space using the fast Fourier transform. By replacing the k-space lines of the Cartesian sampled slices in this sequence, degraded images can be generated with this artefact. The corrupted images using this method are realistic because the mis-triggering artefact results from the misallocations of the k-space lines in the temporal slices in reality. The number of replaced k-space lines is randomly selected from 10 to 20 to obtain different levels of the artefact severity. Figure 2 shows the process of image corruption with this artefact.

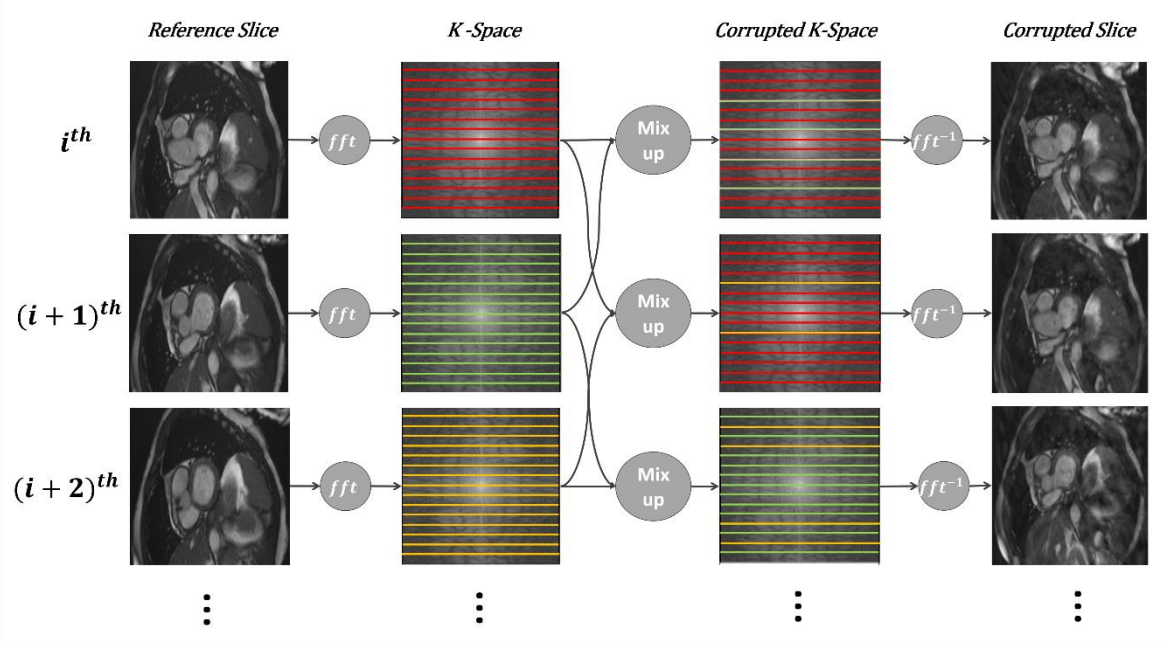


Figure 2. k-space manipulation to generate synthetic cardiac motion artefacts.

2-2-3- Gibbs ringing artefacts

Gibbs ringing or truncation artefacts are due to using Fourier transforms in the reconstruction of the final image. Since to generate an image in magnetic resonance imaging, we have to take a limited sample of the generated signals, so when reconstructing the final image, we have to approximate it using a relatively small number of harmonics in the Fourier representation of the image. Therefore, the truncated Fourier series approximation of a discontinuous signal shows overshoot and ringing near the discontinuities of the primary signal [20]. This case is called the Gibbs phenomenon [21], which causes parallel lines and curves in the final image. Image filtering can be used with an ideal low-pass filter in the frequency domain to add this artefact with different severities to the final image [22]. Thus, the image is transferred to k-space by fast Fourier transform, and then k-space is filtered by an ideal low-pass filter. The radius of the ideal low-pass filter was chosen randomly from 30 to 50 pixels to obtain degraded images with different severities of this artefact. A smaller radius causes a more intense artefact appearance due to the removal of more data from the k-space. The procedure for the k-space manipulation to generate corrupted images is shown in Figure 3.

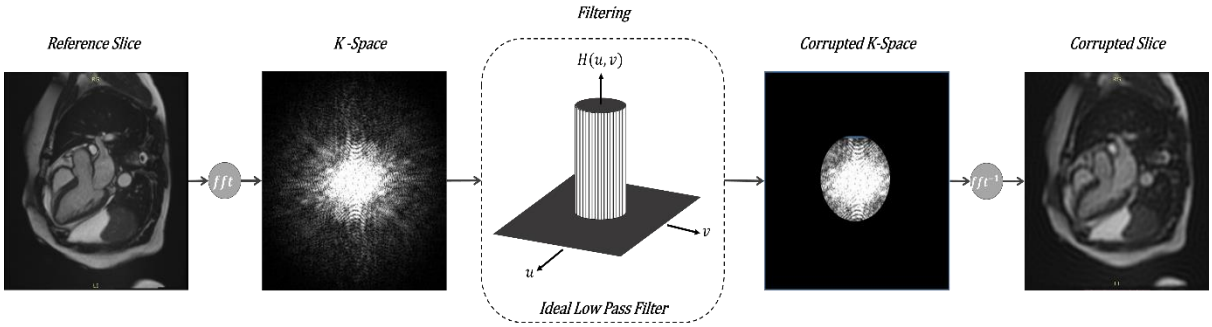


Figure 3. k-space manipulation to generate synthetic Gibbs ringing artefacts.

2-2-4- Aliasing artefacts

Aliasing or wrap-around artefact occurs when the field of view (FOV) is smaller than the area being imaged. Here, the parts outside the FOV are reflected on the other side of the image. It is enough to undersample the k-space to add this artefact to the images by manipulating the k-space. Because the k-space of the images used is Cartesian sampled, the corrupted image can be

obtained by alternately deleting rows or columns of k-space corresponding to an image [23]. An example of image degradation in this way is shown in Figure 4.

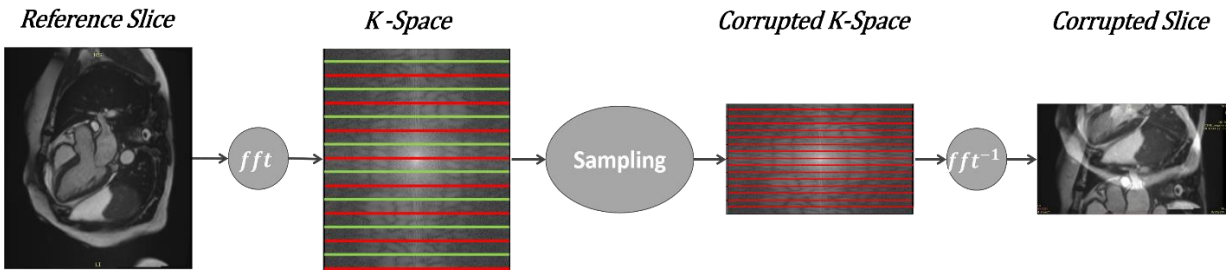


Figure 4. k-space manipulation to generate synthetic aliasing artefact.

2-3- Deep learning models for automated CMR image quality assessment

In this section, we propose a deep-learning model for automatic CMR image quality assessment. First, this model is proposed to identify the artefacts in the spatial domain when data labels are available. The model details for domain adaptation are presented when we use the trained model to test a new dataset. Finally, the proposed model in the frequency domain is described to directly use the model on the k-space generated by the Fourier transform of the corresponding magnitude images.

The proposed model evaluates the quality of each of the 2D slices of a 3D volume of CMR images. The reason for presenting the 2D model is that the artefacts studied in this research have features that can be seen in each 2D slice. Although some previous studies have identified motion artefacts with 3D models like [7], the diversity of artefacts in the current research and the need for a general model to distinguish between types of artefacts led us to use the 2D model. Also, since clinicians ultimately analyse CMR images in a 2D view, the present study used a 2D model to identify artefacts.

2-3-1- Spatial domain

We use a deep learning model to identify artefacts in the spatial domain. This model generally has three essential parts: (1) Feature extraction from CMR images (2) Classification of the image artefacts (3) Domain label predictor for domain adaptation. Figure 5 shows the proposed model in the spatial domain.

The feature extraction section receives the input images with various views, including two, three and four-chamber long-axis and short-axis views. The dimensions of the input images are resized to 90×90 to control the computational overhead. The image resizing operation is done using the Scikit-image Python package to convert the size of all the images to the required size. Four 2D convolutional layers with the details specified in Figure 5 are considered for image feature extraction. Each pair of convolutional layers is followed by a max-pooling layer and a dropout layer with a probability of 0.25. Finally, a flatten followed by a dense layer are at the end of this section, which presents the extracted features in a 512 vector.

The feature vector generated by the feature extractor part is received as input in the label classification section. This section follows a dropout layer with a probability of 0.5 by two dense layers responsible for identifying the artefact labels. Five labels related to respiratory and cardiac motions, aliasing, Gibbs ringing and without artefacts are determined by five terminal neurons. If training data labels are available and supervised or learning from limited annotated training data is on the agenda, we will only use the feature extraction and label classification parts. Otherwise, it is also necessary to use the domain label predictor part for the unsupervised domain adaptation process.

Domain adaptation refers to the concept of learning a discriminative classifier when there is a shift in distributing training and testing data distribution [24]. Domain adaptation approaches are classified into two categories, unsupervised and semi-supervised, based on the possibility of accessing the data labels of the target dataset. The former is when the target dataset is unlabelled, and the latter is when a few samples are labelled. These approaches allow the mapping between the source and target domains to be learned when the target domain is unlabelled or partially labelled. It means that the classifier can be trained on a dataset as the source domain and then used to test the target domain [25]. The main idea of domain adaptation is inspired by [26] and customised for the current study. All three subnets required for feature extraction, artefact classification, and domain label prediction have been redesigned and implemented to meet the objectives of this study.

In the situation where there are two datasets, source and target, where there is a domain shift between them, we should also use the domain classification part of the model for unsupervised domain adaptation. Consider that data with the artefact-type label is available for the source set, and no artefact-type label exists for the target data set. For the domain label predictor part, we assign an additional label representing the source or target domain type to the data of each set. For example, in addition to the artefact-type label, we assign a 0 label to all source data, which indicates that this data belongs to the source set. The data of the target set, which does not have an artefact-type label, only gets a label such as 1, which means that it is the data of the target set. The domain classification part of the model follows a gradient reversal layer (GRL) by three dense layers with the details specified in Figure 5. The GRL layer acts as an identical transformation in forwarding propagation while multiplying the gradient received from the dense layers by a negative value in backpropagation. The reason for using this layer is there are conflicting goals between the parts. If there are two datasets, one as a source with access to the data labels and the target without data labels, the feature extraction part must extract domain-invariant features between different datasets at training time. The source data must have a domain label and an artefact label. In contrast, the target data have the domain label, which differs from the source domain label, to achieve these features. Therefore, the feature extraction part maximises the domain classifier loss and minimises the artefact label classifier loss. Maximising the domain classifier loss makes the distribution of features obtained from the two domains more similar.

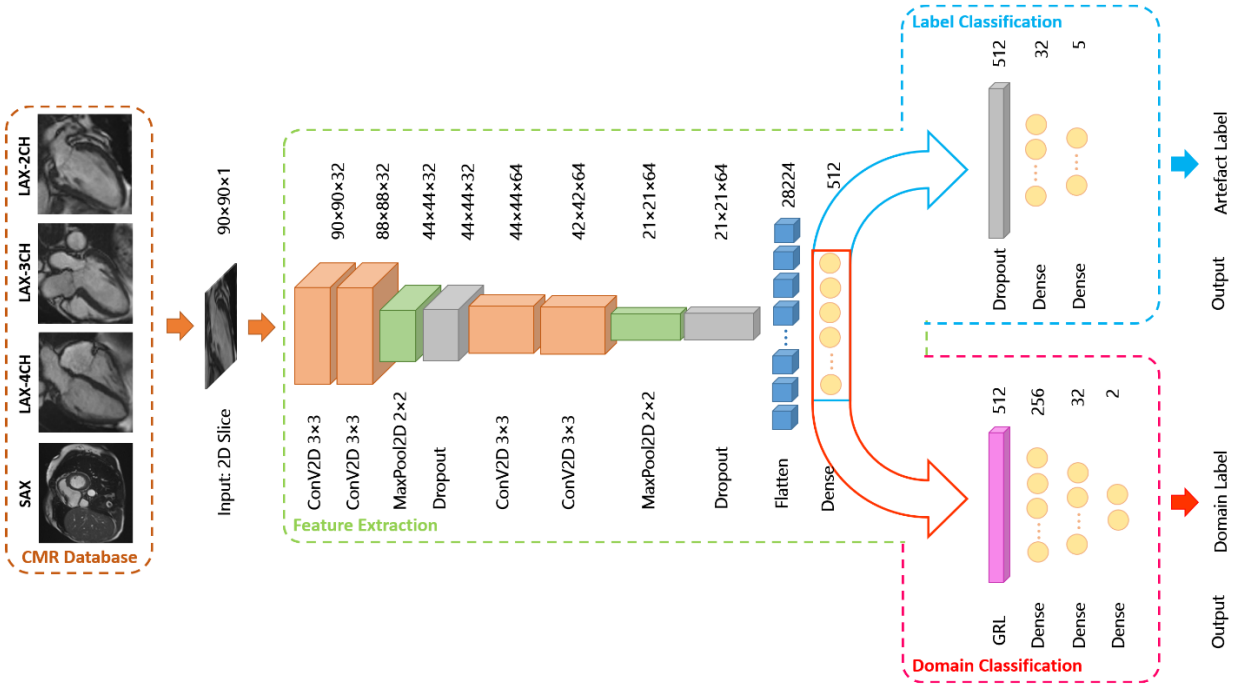


Figure 5. The proposed cardiovascular magnetic resonance image quality assessment model in the spatial domain.

2-3-2- Frequency domain

Two main concerns led us to provide a model in the frequency domain for CMR image quality assessment: (1) the possibility of using the proposed method in the frequency domain on the raw data of k-space before image formation, (2) Increasing the speed of training and testing operations when the model may be used for image quality assessment in large datasets such as UK Biobank. The proposed model is shown in Figure 6.

The proposed model can be used well when working on large datasets and analysing k-space data. Instead of using the ordinary convolutional layer, the proposed model uses Fourier-based CNNs (FCNNs) [27], which dramatically reduces the computational costs of feature extraction. The convolution operation in the frequency domain can be changed to a pointwise multiplication in this model by considering Equation (1).

$$DFT\big(f(x,y) * h(x,y)\big) = DFT(f(x,y)) \cdot DFT(h(x,y)) \quad (1)$$

where DFT represents the discrete Fourier transform, $f(x,y)$ represents the input image, and $h(x,y)$ indicated the filter.

The k-space corresponding to the input image is pointwise multiplied in three filters filled with random complex numbers for feature extraction. This operation is repeated twice, and the size of the filters is the same as the size of the k-space. The output of these steps is given to the max-pooling layer, and the dropout with a probability of 0.25, a flatten, and a dense layer follows this layer. Finally, a 512 vector is transferred to the label classification section; this section's details are shown in Figure 6.

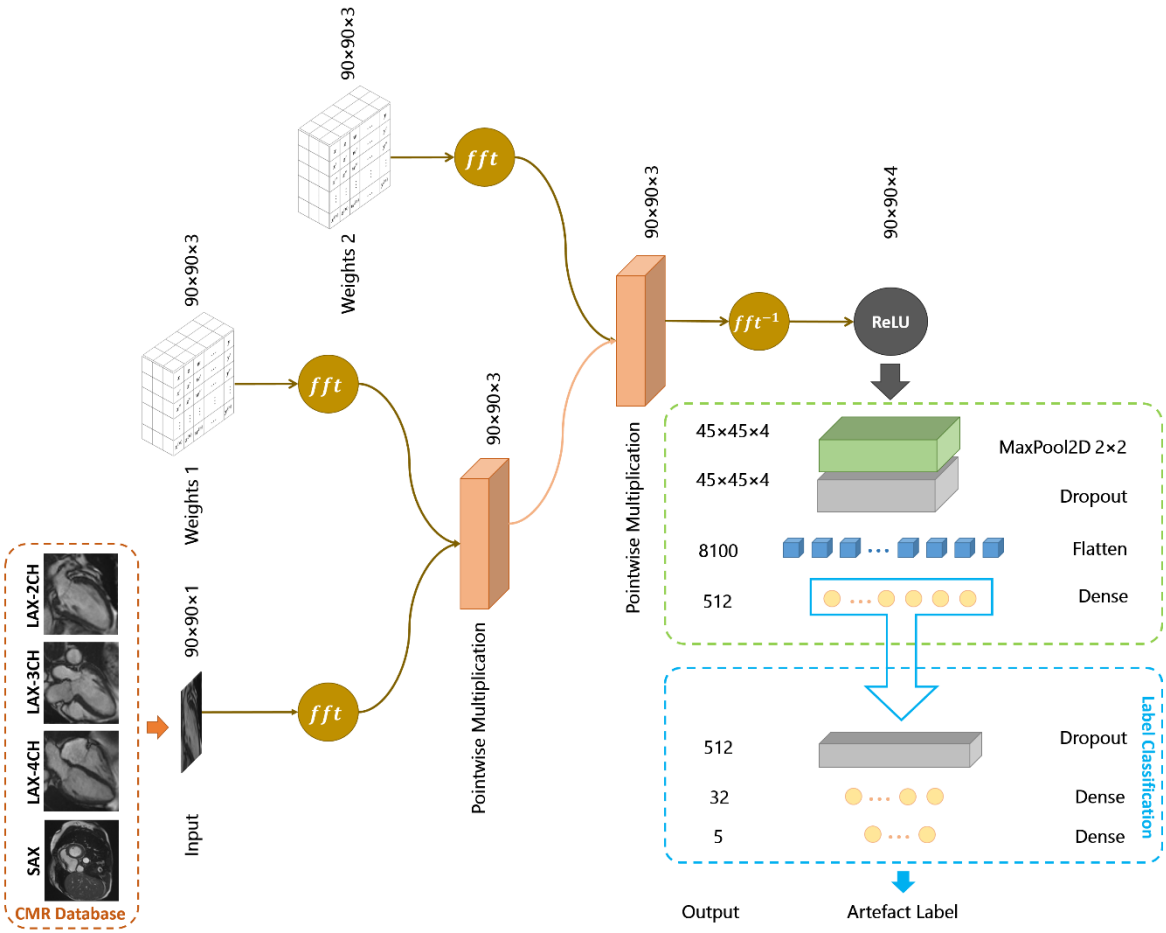


Figure 6. The proposed cardiovascular magnetic resonance image quality assessment model in the frequency domain.

2-4- Model training and Evaluation design

The training process of the existing models was performed in several steps according to different defined experiments. Adam optimiser [28] and cross-entropy loss function [29] have been used to train the models in spatial and frequency domains. The learning rate was set to 1×10^{-3}. Also, the training process has been carried out in 25 epochs and with batch size 128. The activation function for all layers except the last one is rectified linear unit (ReLU). This function is Softmax for the last layer. To validate the model training process and increase the reliability of the results, experiments were performed using 4-fold cross-validation. The best architecture suitable for the current study, including layer placements and selection of hyperparameters, was made based on ablation studies. The hardware platform is a computer equipped with an Intel Core i-7-6700, 16GB of RAM and an NVIDIA GeForce GTX TITAN X GPU. After the article is published, the developed Python code will be made available upon request.

We performed several experiments to evaluate the accuracy and capability of the proposed methods. In the first group of these experiments, the proposed deep learning models in spatial and frequency domains are learned in a supervised learning manner. This way, 75% of the data is a training set, a part of this 75% is the validation set, and the rest 25% is for testing. Given one of the most fundamental challenges, namely the lack of access to labelled data, the second group of experiments is assumed to be supervised using limited training data. In this group of experiments, 25% of the data is used for training and the rest for testing. These experiments make the results comparable in terms of the effectiveness of the amount of annotated data used in the training process. The third set of experiments is performed to evaluate the proposed unsupervised domain adaptation model. All parts of the Figure 5 model are used to perform these experiments. In these experiments, the total number of images of the UK Biobank, YU, CMR-Tehran, and UCIII datasets after adding artefacts and data augmentation are 255775, 106026, 17901 and 4784, respectively. To learn the model, any sample of the source set has both the domain and artefact label, but instances of the target set have only the domain label. UK Biobank data is first considered a source set in these experiments, and other data sets are considered a target. The same experiments are then repeated by considering the YU dataset as the source. The difference between training and testing speed and the accuracy in the spatial and frequency domains are also evaluated. These experiments examine the capability of the proposed models in the CMR image quality assessment. The training and testing data for all five classes are considered almost balanced to make a fair comparison and avoid bias.

Some experiments were also conducted on the real data to check the performance of the proposed models. These experiments can show how similar the synthetic corrupted images are to the real degraded images. For this purpose, UCIII and CMR-Tehran datasets were first labelled by a radiologist with three years of experience in cardiac imaging from the perspective of the artefacts examined in this study. Then, these labels were confirmed by a cardiologist with 15 years of experience in cardiac imaging. The data used to evaluate the frequency domain model in these experiments are obtained from the Fourier transform of the corresponding magnitude images and are not scanner-acquired k-spaces.

Five metrics, including accuracy (ACC), precision (PR), recall (RE), F-measure, and area under the ROC curve (AUC), given in Equations 2 to 6, are used to evaluate the results of the proposed models quantitatively.

$$ACC = \frac{TP + TN}{TP + FP + TN + FN} \quad (2)$$

$$PR = \frac{TP}{TP + FP} \quad (3)$$

$$RE = \frac{TP}{TP + FN} \quad (4)$$

$$F - measure = \frac{2 \times PR \times RE}{PR + RE} \quad (5)$$

$$AUC = \int_0^1 \Pr[TP]\,(v)\,dv \quad (6)$$

where TP, FP, TN, and FN indicate true positive, false positive, true negative, and false negative. AUC denotes the overall success of an experiment where $\Pr[TP]$ is a function of $v = \Pr[FP]$.

3- Results

3-1- Results of Supervised learning experiments

In the first experiment, all datasets, including reference images and corrupted images generated by k-space manipulation methods, were used to train and test the models in a supervised manner. After the image corruption by k-space manipulating techniques and data augmentation, 255775, 144411, 23638 and 6080 CMR images were prepared for UK Biobank, YU, CMR-Tehran, and UCIII datasets, respectively. The results to assess the quality of CMR images in a supervised manner for respiratory and cardiac motions, Gibbs ringing and aliasing artefacts are given in Table 1. Besides, the comparison of the models' performance for different views of the images is given in Table S2 (See supplementary material).

Table 1. Results for supervised evaluation of the proposed CMR image quality assessment models are based on accuracy, precision, recall, F-measure and AUC metrics in spatial and frequency domains. Results are based on 4-fold cross-validation ($Mean \pm std$).

Dataset \ Metrics		Accuracy	Precision	Recall	F-measure	AUC
UK Biobank	Spatial	$99 \bullet 40 \pm 0 \bullet 09$	$99 \bullet 41 \pm 0 \bullet 09$	$99 \bullet 40 \pm 0 \bullet 09$	$99 \bullet 40 \pm 0 \bullet 09$	$99 \bullet 89 \pm 0 \bullet 07$
	Frequency	$87 \bullet 46 \pm 0 \bullet 61$	$87 \bullet 59 \pm 0 \bullet 64$	$87 \bullet 31 \pm 0 \bullet 57$	$87 \bullet 45 \pm 0 \bullet 61$	$98 \bullet 64 \pm 0 \bullet 19$
YU	Spatial	$75 \bullet 78 \pm 9 \bullet 17$	$76 \bullet 88 \pm 9 \bullet 27$	$75 \bullet 04 \pm 9 \bullet 06$	$75 \bullet 95 \pm 9 \bullet 15$	$91 \bullet 26 \pm 5 \bullet 44$
	Frequency	$63 \bullet 76 \pm 12 \bullet 35$	$67 \bullet 94 \pm 12 \bullet 61$	$58 \bullet 23 \pm 13 \bullet 02$	$62 \bullet 70 \pm 12 \bullet 92$	$88 \bullet 47 \pm 7 \bullet 71$
CMR-Tehran	Spatial	$67 \bullet 87 \pm 4 \bullet 10$	$68 \bullet 46 \pm 3 \bullet 97$	$67 \bullet 38 \pm 4 \bullet 13$	$67 \bullet 91 \pm 4 \bullet 05$	$87 \bullet 23 \pm 2 \bullet 32$
	Frequency	$58 \bullet 48 \pm 0 \bullet 41$	$62 \bullet 83 \pm 0 \bullet 65$	$48 \bullet 37 \pm 3 \bullet 78$	$54 \bullet 60 \pm 2 \bullet 46$	$88 \bullet 63 \pm 0 \bullet 59$
UCIII	Spatial	$89 \bullet 46 \pm 1 \bullet 01$	$89 \bullet 61 \pm 1 \bullet 20$	$89 \bullet 03 \pm 0 \bullet 62$	$89 \bullet 32 \pm 0 \bullet 87$	$97 \bullet 10 \pm 0 \bullet 51$
	Frequency	$80 \bullet 25 \pm 0 \bullet 94$	$83 \bullet 91 \pm 1 \bullet 16$	$76 \bullet 34 \pm 2 \bullet 00$	$79 \bullet 94 \pm 1 \bullet 50$	$96 \bullet 83 \pm 0 \bullet 34$

The results of the second group of experiments are given in Table 2 for spatial and frequency domains when the training data is limited.

Table 2. Evaluation of learning from limited training data using the proposed CMR image quality assessment models is based on accuracy, precision, recall, F-measure and AUC metrics in spatial and frequency domains. Results are based on 4-fold cross-validation ($Mean \pm std$).

Dataset \ Metrics		Accuracy	Precision	Recall	F-measure	AUC
UK Biobank	Spatial	$96 \bullet 37 \pm 0 \bullet 66$	$96 \bullet 37 \pm 0 \bullet 66$	$96 \bullet 37 \pm 0 \bullet 66$	$96 \bullet 37 \pm 0 \bullet 66$	$99 \bullet 07 \pm 0 \bullet 13$
	Frequency	$79 \bullet 78 \pm 1 \bullet 56$	$80 \bullet 36 \pm 1 \bullet 58$	$79 \bullet 23 \pm 1 \bullet 52$	$79 \bullet 79 \pm 1 \bullet 54$	$96 \bullet 45 \pm 0 \bullet 34$
YU	Spatial	$53 \bullet 25 \pm 6 \bullet 06$	$53 \bullet 82 \pm 6 \bullet 11$	$52 \bullet 43 \pm 6 \bullet 23$	$53 \bullet 11 \pm 6 \bullet 17$	$77 \bullet 68 \pm 4 \bullet 75$
	Frequency	$45 \bullet 66 \pm 4 \bullet 42$	$49 \bullet 86 \pm 6 \bullet 45$	$37 \bullet 54 \pm 4 \bullet 28$	$42 \bullet 67 \pm 4 \bullet 19$	$77 \bullet 61 \pm 5 \bullet 68$
CMR-Tehran	Spatial	$54 \bullet 04 \pm 4 \bullet 64$	$54 \bullet 30 \pm 4 \bullet 92$	$52 \bullet 36 \pm 6 \bullet 54$	$53 \bullet 30 \pm 5 \bullet 78$	$79 \bullet 20 \pm 2 \bullet 61$
	Frequency	$45 \bullet 96 \pm 2 \bullet 47$	$49 \bullet 06 \pm 4 \bullet 30$	$36 \bullet 63 \pm 1 \bullet 93$	$41 \bullet 92 \pm 2 \bullet 62$	$81 \bullet 54 \pm 1 \bullet 81$
UCIII	Spatial	$79 \bullet 35 \pm 6 \bullet 22$	$79 \bullet 75 \pm 5 \bullet 92$	$78 \bullet 92 \pm 6 \bullet 57$	$79 \bullet 33 \pm 6 \bullet 24$	$92 \bullet 23 \pm 2 \bullet 69$
	Frequency	$59 \bullet 40 \pm 4 \bullet 65$	$66 \bullet 98 \pm 6 \bullet 26$	$50 \bullet 94 \pm 3 \bullet 33$	$57 \bullet 83 \pm 4 \bullet 25$	$86 \bullet 26 \pm 3 \bullet 44$

3-2- Results of Unsupervised domain adaptation model

This section presents the results of experiments designed to evaluate the proposed domain adaptation method. To compare the results, three modes of implementation are considered. The highest accuracy is achieved when the model can receive artefact labels related to the samples of the target set during training (Upper band). The least accuracy also occurs when the model gets no instances of the target set at the time of training, and the trained model is then evaluated on the samples of the target set (Lower band). In the first experiment, we consider a scenario where the UK Biobank dataset is a source set, and the other datasets are considered a target. The results of the first experiment are shown in Table 3.

Table 3. Results of the first unsupervised domain adaptation experiment (Source: UK Biobank, Target: YU, CMR-Tehran and UCIII) based on accuracy metric (in the range of 0 to 100). The numbers in parentheses in the row corresponding to the results of the proposed method indicate the gap coverage between the upper and lower bands.

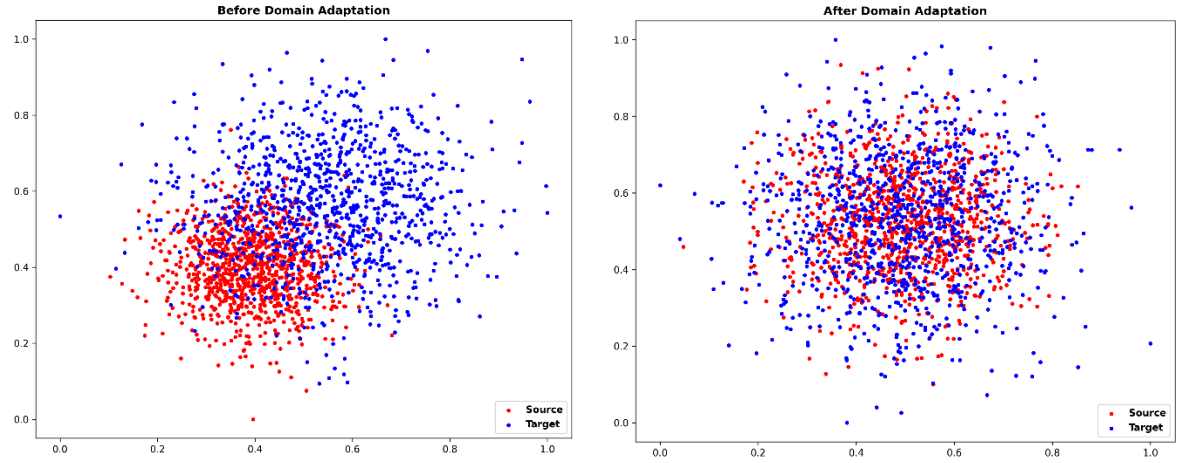
Training mode \ Target set	YU	CMR-Tehran	UCIII
Source only	26.58	15.90	30.50
Proposed method	32.44 (+11.91%)	24.51 (+16.57%)	32.82 (+3.93%)
Train on the annotated target set	75.78	67.87	89.46

In the second experiment of this section, the results are presented in a situation where the source set is YU, and the target sets are the other datasets. The results are given in Table 4.

Table 4. Results of the second unsupervised domain adaptation experiment (Source: YU, Target: UK Biobank, CMR-Tehran and UCIII) based on accuracy metric (in the range of 0 to 100). The numbers in parentheses in the row corresponding to the results of the proposed method indicate the gap coverage between the upper and lower bands.

Training mode \ Target set	UK Biobank	CMR-Tehran	UCIII
Source only	28.62	50.53	31.80
Proposed method	30.17 (+2.19%)	43.88 (-38.35 %)	41.52 (+16.86%)
Train on the annotated target set	99.41	67.87	89.46

Figure 7 shows the effect of the proposed method on the distribution of 512 feature vectors of the last feature extraction layer before and after domain adaptation. The feature vectors of 1000 images are visualised in this figure using the T-distributed stochastic neighbour embedding (t-SNE) method [30]. In this experiment, the YU dataset is the source set (red points), and the UCIII dataset is the target set (blue points). The proposed domain adaptation method could bring the two distributions of feature vectors much closer after adaptation.



(a) (b)

Figure 7. The effect of domain adaptation on changing the distribution of feature values extracted for 1000 images. The feature vectors extracted by the first part of the proposed model in the spatial domain are visualised using t-SNE (a) before and (b) after domain adaptation. The red and blue points correspond to the YU and UCIII datasets, respectively.

3-3- Spatial versus frequency domain

The time required for supervised training and testing the models in the spatial and frequency domains is compared in Table 5. These times are measured in seconds. The model in the frequency domain can be learned and tested more quickly. However, this increase in training and testing speed comes with decreased accuracy in this domain. Therefore, we investigated increasing the number of training samples in another experiment to improve the frequency domain's accuracy. The results show that increasing training samples can improve accuracy while training time increases slightly. Random subsets from the UK Biobank dataset were used to perform these comparisons. The results are shown in Table 6 and Figure 8.

Table 5. Comparison of supervised training and test time differences in proposed deep learning models in spatial and frequency domains for each dataset. The times mentioned are measured in seconds ($Mean \pm std$) based on 4-fold cross-validation. Speed up ratio expresses the increase in training speed in the frequency domain model compared to the spatial domain model.

Dataset \ Domain		Spatial (sec)	Frequency (sec)	Speed up ratio
UK Biobank	Train	2249 • 44 ± 3 • 30	258 • 50 ± 1 • 93	8.70
	Test	32 • 56 ± 0 • 26	7 • 44 ± 2 • 22	4.38
YU	Train	1343 • 90 ± 0 • 83	156 • 39 ± 0 • 17	8.59
	Test	10 • 06 ± 0 • 28	2 • 69 ± 0 • 91	3.74
CMR-Tehran	Train	277 • 71 ± 0 • 87	52 • 26 ± 0 • 50	5.31
	Test	2 • 91 ± 0 • 19	0 • 33 ± 0 • 03	8.82
UCIII	Train	133 • 38 ± 1 • 04	42 • 10 ± 0 • 52	3.17
	Test	1 • 79 ± 0 • 06	0 • 09 ± 0 • 01	19.89

Table 6. Comparison of accuracy, precision and recall based on the different number of training samples in spatial and frequency domains and its effect on training time.

	# of training samples	# of model parameters	Accuracy	Precision	Recall	Time (Sec)
Spatial	30,125	≃ 14.6 M	89.56%	89.63%	89.47%	394.80
Frequency (Run 1)	30,125	≃ 4.3 M	76.24%	76.33%	75.92%	43.22
Frequency (Run 2)	180,750	≃ 4.3 M	87.99%	88.12%	87.83%	255.04

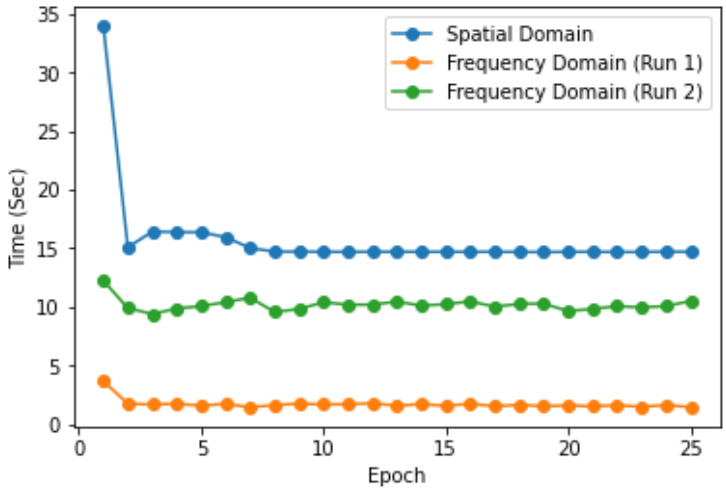


Figure 8. Training time chart in spatial and frequency domains.

3-4- Real versus synthetic corrupted images

The evaluation results of the proposed models in the spatial and frequency domains for experiments related to training from limited data on the real and synthetic corrupted images are shown in Table 7. Besides, the evaluation results of the proposed domain adaptation method, where the source sets are the UK Biobank and the YU dataset, and the target sets are the CMR-Tehran, and UCIII, are tabulated in Table 8. The results indicate that although there is a slight drop in the accuracy obtained on the real data, the achieved results are close.

Table 7. Evaluation of learning from limited training data using the proposed CMR image quality assessment models based on accuracy metric in spatial and frequency domains. Results were obtained using synthetic and real data based on 4-fold cross-validation ($Mean \pm std$).

Dataset \ Type of Data		Synthetic Data	Real Data
CMR-Tehran	Spatial	$54 \bullet 04 \pm 4 \bullet 64$	$53 \bullet 54 \pm 3 \bullet 87$
	Frequency	$45 \bullet 96 \pm 2 \bullet 47$	$45 \bullet 58 \pm 2 \bullet 19$
UCIII	Spatial	$79 \bullet 35 \pm 6 \bullet 22$	$75 \bullet 44 \pm 5 \bullet 33$
	Frequency	$59 \bullet 40 \pm 4 \bullet 65$	$57 \bullet 03 \pm 4 \bullet 25$

Table 8. Evaluation of the unsupervised domain adaptation using synthetic and real data based on accuracy metric (in the range of 0 to 100).

Dataset \ Type of Data		Synthetic Data	Real Data
Source: UK Biobank	**Target: CMR-Tehran**	24.51	22.07
	Target: UCIII	32.82	25.83
Source: YU	**Target: CMR-Tehran**	43.88	43.19
	Target: UCIII	41.52	36.70

4- Discussion

In this study, we aim to address CMR image quality assessment. For this purpose, we tried to propose a comprehensive model for identifying and reporting common artefacts of these images. Respiratory motion, cardiac motion, Gibbs ringing and aliasing artefacts were examined in this study, which is superior to the previous studies regarding the number of artefacts examined. Input images with different views were used to provide a model for recognising these artefacts. Methods were also proposed to add artefacts to CMR images by manipulating the k-space to generate synthetic corrupted images. In addition, detecting artefacts in spatial and frequency domains was investigated to assess the quality using k-space and with no image reconstruction. Therefore, deep learning models were proposed in both spatial and frequency domains. A deep CNN architecture in the spatial domain and a deep Fourier-based CNN architecture in the frequency domain was used.

Several experiments were designed and performed to evaluate the proposed models. In the first series of experiments, the models were trained on the prepared datasets in a supervised manner. These experiments' results show that the proposed deep learning models can detect the mentioned CMR artefacts. The results show a high AUC for the proposed models, indicating the models' suitability to assess the quality of CMR images. The optimal architecture of models in terms of the number of free parameters or weights and other hyperparameters was selected by ablation studies. In these ablation studies, by increasing and decreasing the number of free parameters, a state of the models was found in which the models have the best performance. Although this process is time-consuming, it is used to find the best possible neural network architecture. Besides, The learning curves related to the

supervised training of the proposed models in the spatial and frequency domains on the UK Biobank dataset are shown in Figure S3 (See supplementary material).

One of the main challenges and limitations in deep network training is accessing large amounts of labelled data. Besides, many of the presented deep learning models may provide satisfactory results on the studied dataset, but their use on other datasets with different distributions may not provide promising results. Thus, the concept of domain adaptation was considered to overcome these challenges. We tried to select and augment datasets with considerable distribution gaps in this study. Due to the high quality of UK Biobank images, an attempt was made to choose other datasets of different distribution, imaging parameters, subjects being imaged, and image quality. The three datasets used to evaluate the domain adaptation approach have a vast domain shift compared to the UK Biobank dataset.

The proposed domain adaptation method showed a slight improvement when no labels were available for the target datasets. The reason can be attributed to the large domain shifts in the CMR imaging datasets. Although the improvement obtained by the proposed domain adaptation method seems insignificant, it can provide a good motivation for research in this field, given the importance of the problem. Besides, it was also found in one of the experiments on the proposed domain adaptation model that using this method has failed to succeed. A similar lack of success has also been reported in the existing studies in this field [26, 31]. The domain shift in CMR datasets is much wider than the examples in the cited articles due to different content, imaging sequences, and acquisition devices. Therefore, the accuracy for CMR images was expected to be lower than that of, for example, MNIST dataset images.

Another achieved goal of this study was to increase the speed of deep learning-based models. For this purpose, a Fourier-based convolutional model was proposed. Based on the analysis of the results, these networks could perform well while increasing training and testing speed. Using the model in the frequency domain eliminated the need to define fixed-size kernels in convolution operations. The reason is that the kernel size was considered the same size as the input image. Considering kernels of the same size as the image would also extract global features from the images. Besides, there was no need to move the sliding window over the entire image, which is time-consuming. Large-size images can also be used in these models because the complexity of feature extraction operations is much less than the spatial domain. According to the experiments performed in this study, the desired accuracy can be achieved in a shorter time by increasing the number of samples.

The datasets used in this study include magnitude images, which means no scanner-acquired k-space data from multi-channel receiver coils were available. So creating k-space using Fourier transform from these data can only mimic k-space due to the lack of access to phase data. This issue is the most critical limitation of the study. Also, some cases that may cause image corruption during imaging cannot be simulated with the proposed methods. For example, only bulk motions can be simulated with these techniques, and deformable motions cannot be implemented. Therefore, although the proposed methods for adding artefacts to images do not exactly correspond to reality, they can simulate something quasi-real.

Considering that in the current study, k-space data from multi-channel receiver coils were unavailable, the frequency domain model's behaviour is unknown in real-time applications for image quality assessment. However, the results obtained on the k-space achieved from the Fourier transform of the magnitude images showed that the proposed frequency space model has the potential for real-time applications. Of course, this issue requires more studies on the scanner-acquired k-space data. The cost of image reconstruction using the Fourier transform is avoided, and the quality can be assessed immediately after filling the k-space, assuming the efficiency of the frequency model on the original k-space data.

CMR image quality control involves examining the full heart coverage, signal-to-noise ratio, and imaging artefacts. A different approach is taken to explore each of these cases. In some studies [8-10], whole heart or left ventricle coverage was examined. These studies investigated short-axis cine CMR volumes for the presence or absence of apex and basal slices. Other researches [5-7] have focused on CMR imaging artefacts. The proposed models of those studies are 3D models based on the evaluation of artefacts in a 3D CMR volume. Since physicians ultimately examine CMR images in a 2D view, the current study used a 2D model to identify the artefacts.

Evaluations were also done on data with real artefacts to check the effectiveness of the proposed models for clinical applications. The results from applying the models to the data with real artefacts showed a slight decrease in the obtained metrics. However, the models' performance on these data was almost similar to the performance on the synthetic degraded data. An important issue that should be considered in analysing the results is that when the number of data used to train the models has increased, the proposed models showed better performance and more stability due to the data-hungry nature of deep learning models. The values of the obtained metrics are reduced in the other datasets due to their smaller volume, and in the frequency model, the recall is also reduced compared to the precision. Therefore, although we used different datasets to show the efficiency of the proposed models, the larger training set led to better model stability. Of course, other factors such as spatial resolution, imaging parameters, amount of image degradation when adding artefacts to reference images (which are considered random for diversity in this study), and image contents (the UCIII dataset consists of small animal CMR images) can also affect the performance of the models on YU, UCIII, and CMR-Tehran datasets.

5- Conclusion

This study proposes two models in the spatial and frequency domains for CMR image quality assessment. The proposed models can recognise the four common artefacts, including respiratory motion, cardiac motion, Gibbs ringing and aliasing, by receiving the input image or k-space. Besides, the problem of lack of access to annotated data was addressed by proposing a domain adaptation method. The analysis of the results obtained on four different datasets shows that the proposed models can properly detect the mentioned artefacts. Despite the poor performance of the proposed domain adaptation method, its results can be used as a baseline for further studies. Presenting more comprehensive models covering more artefacts and using new methods such as meta-learning will be among the purposes of the future study. Developing a hybrid model that takes advantage of both spatial

and frequency domains for CMR image quality assessment and the proposed frequency domain model evaluation on the scanner-acquired k-space data are also potential future works.

Acknowledgements
This research has been done using the UK Biobank database under Application 11350.

Conflict of Interest Statement
There are no conflicts of interest to declare.

References

[1] W. P. Bandettini and A. E. Arai, "Advances in clinical applications of cardiovascular magnetic resonance imaging," *Heart,* vol. 94, no. 11, pp. 1485-1495, 2008.
[2] A. Kumar, D. J. Patton, and M. G. Friedrich, "The emerging clinical role of cardiovascular magnetic resonance imaging," *Canadian Journal of Cardiology,* vol. 26, no. 6, pp. 313-322, 2010.
[3] M. Salerno *et al.*, "Recent advances in cardiovascular magnetic resonance: techniques and applications," *Circulation: Cardiovascular Imaging,* vol. 10, no. 6, p. e003951, 2017.
[4] L. S. Chow and R. Paramesran, "Review of medical image quality assessment," *Biomedical signal processing and control,* vol. 27, pp. 145-154, 2016.
[5] G. Tarroni *et al.*, "Learning-based quality control for cardiac MR images," *IEEE transactions on medical imaging,* vol. 38, no. 5, pp. 1127-1138, 2018.
[6] G. Tarroni *et al.*, "Large-scale Quality control of cardiac imaging in population Studies: Application to UK Biobank," *Scientific reports,* vol. 10, no. 1, pp. 1-11, 2020.
[7] I. Oksuz *et al.*, "Automatic CNN-based detection of cardiac MR motion artefacts using k-space data augmentation and curriculum learning," *Medical image analysis,* vol. 55, pp. 136-147, 2019.
[8] L. Zhang *et al.*, "Automatic assessment of full left ventricular coverage in cardiac cine magnetic resonance imaging with fisher-discriminative 3-D CNN," *IEEE Transactions on Biomedical Engineering,* vol. 66, no. 7, pp. 1975-1986, 2018.
[9] L. Zhang, M. Pereañez, S. K. Piechnik, S. Neubauer, S. E. Petersen, and A. F. Frangi, "Multi-input and dataset-invariant adversarial learning (MDAL) for left and right-ventricular coverage estimation in cardiac MRI," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2018: Springer, pp. 481-489.
[10] L. Zhang, A. Gooya, and A. F. Frangi, "Semi-supervised assessment of incomplete LV coverage in cardiac MRI using generative adversarial nets," in *International Workshop on Simulation and Synthesis in Medical Imaging*, 2017: Springer, pp. 61-68.
[11] M. Osadebey, M. Pedersen, D. Arnold, and K. Wendel-Mitoraj, "Image quality evaluation in clinical research: A case study on brain and cardiac mri images in multi-center clinical trials," *IEEE journal of translational engineering in health and medicine,* vol. 6, pp. 1-15, 2018.
[12] S. Nabavi, M. Hashemi, M. Ebrahimi Moghaddam, A. A. Abin, and A. F. Frangi, "Automated cardiac coverage assessment in cardiovascular magnetic resonance imaging using an explainable recurrent 3D dual-domain convolutional network," *Medical Physics,* vol. 51, no. 12, pp. 8789-8803, 2024.
[13] S. Nabavi, H. Simchi, M. E. Moghaddam, A. A. Abin, and A. F. Frangi, "A generalised deep meta-learning model for automated quality control of cardiovascular magnetic resonance images," *Computer Methods and Programs in Biomedicine,* vol. 242, p. 107770, 2023.
[14] S. E. Petersen *et al.*, "UK Biobank's cardiovascular magnetic resonance protocol," *Journal of cardiovascular magnetic resonance,* vol. 18, no. 1, pp. 1-7, 2015.
[15] A. Andreopoulos and J. K. Tsotsos, "Efficient and generalizable statistical models of shape and appearance for analysis of cardiac MRI," *Medical Image Analysis,* vol. 12, no. 3, pp. 335-357, 2008.
[16] J. F. Abascal, P. Montesinos, E. Marinetto, J. Pascau, and M. Desco, "Comparison of total variation with a motion estimation based compressed sensing approach for self-gated cardiac cine MRI in small animal studies," *PloS one,* vol. 9, no. 10, p. e110594, 2014.
[17] T. Budrys, V. Veikutis, S. Lukosevicius, R. Gleizniene, E. Monastyreckiene, and I. Kulakiene, "Artifacts in magnetic resonance imaging: how it can really affect diagnostic image quality and confuse clinical diagnosis?," *Journal of Vibroengineering,* vol. 20, no. 2, pp. 1202-1213, 2018.
[18] P. F. Ferreira, P. D. Gatehouse, R. H. Mohiaddin, and D. N. Firmin, "Cardiovascular magnetic resonance artefacts," *Journal of Cardiovascular Magnetic Resonance,* vol. 15, no. 1, pp. 1-39, 2013.
[19] B. Lorch, G. Vaillant, C. Baumgartner, W. Bai, D. Rueckert, and A. Maier, "Automated detection of motion artefacts in MR imaging using decision forests," *Journal of medical engineering,* vol. 2017, 2017.
[20] L. F. Czervionke, J. M. Czervionke, D. L. Daniels, and V. M. Haughton, "Characteristic features of MR truncation artifacts," *American journal of roentgenology,* vol. 151, no. 6, pp. 1219-1228, 1988.
[21] J. W. Gibbs, "Fourier's series," *Nature,* vol. 1539, no. 59, p. 606, 1899.
[22] A. C. Bovik and S. T. Acton, "Basic linear filtering with application to image enhancement," in *The Essential Guide to Image Processing*: Elsevier, 2009, pp. 225-239.
[23] S. Y. Huang, R. T. Seethamraju, P. Patel, P. F. Hahn, J. E. Kirsch, and A. R. Guimaraes, "Body MR imaging: artifacts, k-Space, and solutions," *Radiographics,* vol. 35, no. 5, pp. 1439-1460, 2015.

[24] A. Farahani, S. Voghoei, K. Rasheed, and H. R. Arabnia, "A brief review of domain adaptation," *Advances in data science and information engineering,* pp. 877-894, 2021.
[25] W. M. Kouw and M. Loog, "A review of domain adaptation without target labels," *IEEE transactions on pattern analysis and machine intelligence,* vol. 43, no. 3, pp. 766-785, 2019.
[26] Y. Ganin and V. Lempitsky, "Unsupervised domain adaptation by backpropagation," in *International conference on machine learning*, 2015: PMLR, pp. 1180-1189.
[27] H. Pratt, B. Williams, F. Coenen, and Y. Zheng, "Fcnn: Fourier convolutional neural networks," in *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, 2017: Springer, pp. 786-798.
[28] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," presented at the International Conference for Learning Representations (ICLR), May 2015.
[29] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.
[30] L. Van der Maaten and G. Hinton, "Visualizing data using t-SNE," *Journal of machine learning research,* vol. 9, no. 11, 2008.
[31] B. Gholami, P. Sahu, O. Rudovic, K. Bousmalis, and V. Pavlovic, "Unsupervised multi-target domain adaptation: An information theoretic approach," *IEEE Transactions on Image Processing,* vol. 29, pp. 3993-4002, 2020.

Shahabedin Nabavi received his Ph.D. in Artificial Intelligence from Shahid Beheshti University in 2024, with a dissertation on *Automated Cardiovascular MR Image Quality Assessment*. His research interests include **medical image analysis, deep learning, generative models, and quality assessment of cardiovascular imaging.**

Hossein Simchi received his M.Sc. in Artificial Intelligence from the Faculty of Computer Science and Engineering, Shahid Beheshti University. His research interests include **artificial intelligence, machine learning, deep learning, interpretable AI, and healthcare applications**.

Mohsen Ebrahimi Moghaddam is a Professor in the Faculty of Computer Science and Engineering at Shahid Beheshti University, Iran. His research interests include **image processing, computer vision, artificial intelligence, and operating systems**.

Ahmad Ali Abin is an Associate Professor in the Faculty of Computer Science and Engineering at Shahid Beheshti University, Iran. His research interests include **machine learning, deep learning, and forecasting**.

Alejandro F. Frangi is the Bicentennial Turing Chair in Computational Medicine at the University of Manchester, UK, where he also holds a RAEng Chair. He is affiliated with KU Leuven and the Alan Turing Institute. His research interests include **medical image computing, computational medicine, and in-silico trials**.