



# Article Reproducibility and Repeatability of Coronary Computed Tomography Angiography (CCTA) Image Segmentation in Detecting Atherosclerosis: A Radiomics Study

Mardhiyati Mohd Yunus <sup>1,2</sup>, Akmal Sabarudin <sup>1</sup>, Muhammad Khalis Abdul Karim <sup>3,\*</sup>, Puteri N. E. Nohuddin <sup>4,5</sup>, Isa Azzaki Zainal <sup>6</sup>, Mohd Shahril Mohd Shamsul <sup>6</sup> and Ahmad Khairuddin Mohamed Yusof <sup>7</sup>

- <sup>1</sup> Programme of Diagnostic Imaging and Radiotherapy, Faculty of Health Sciences, Universiti Kebangsaan Malaysia (UKM), Kuala Lumpur 56000, Malaysia
- <sup>2</sup> Programme of Medical Imaging, Faculty of Health Sciences, Universiti Selangor (UNISEL), Batang Berjuntai 40000, Selangor, Malaysia
- <sup>3</sup> Department of Physics, Faculty of Science, Universiti Putra Malaysia (UPM), Serdang 43400, Selangor, Malaysia
- <sup>4</sup> Institute of IR4.0, Universiti Kebangsaan Malaysia (UKM), Bangi 43600, Selangor, Malaysia
- <sup>5</sup> Faculty of Business, Higher College of Technology, Sharjah P.O. Box 7947, United Arab Emirates
- <sup>6</sup> Department of Radiology, Hospital Canselor Tunku Muhriz (HCTM), Kuala Lumpur 56000, Malaysia
- Imaging Centre, Institut Jantung Negara (IJN), Kuala Lumpur 50400, Malaysia
- Correspondence: mkhalis@upm.edu.my

7

Abstract: Atherosclerosis is known as the leading factor in heart disease with the highest mortality rate among the Malaysian population. Usually, the gold standard for diagnosing atherosclerosis is by using the coronary computed tomography angiography (CCTA) technique to look for plaque within the coronary artery. However, qualitative diagnosis for noncalcified atherosclerosis is vulnerable to false-positive diagnoses, as well as inconsistent reporting between observers. In this study, we assess the reproducibility and repeatability of segmenting atherosclerotic lesions manually and semiautomatically in CCTA images to identify the most appropriate CCTA image segmentation method for radiomics analysis to quantitatively extract the atherosclerotic lesion. Thirty (30) CCTA images were taken retrospectively from the radiology image database of Hospital Canselor Tuanku Muhriz (HCTM), Kuala Lumpur, Malaysia. We extract 11,700 radiomics features which include the first-order, second-order and shape features from 180 times of image segmentation. The interest vessels were segmentized manually and semiautomatically using LIFEx (Version 7.0.15, Institut Curie, Orsay, France) software by two independent radiology experts, focusing on three main coronary blood vessels. As a result, manual segmentation with a soft-tissuewindowing setting yielded higher repeatability as compared to semiautomatic segmentation with a significant intraclass correlation coefficient (intra-CC) 0.961 for thefirst-order and shape features; intra-CC of 0.924 for thesecond-order features with p < 0.001. Meanwhile, the semiautomatic segmentation has higher reproducibility as compared to manual segmentation with significant interclass correlation coefficient (inter-CC) of 0.920 (first-order features) and a good interclass correlation coefficient of 0.839 for the second-order features with p < 0.001. The first-order, shape order and second-order features for both manual and semiautomatic segmentation have an excellent percentage of reproducibility and repeatability (intra-CC > 0.9). In conclusion, semi-automated segmentation is recommended for inter-observer study while manual segmentation with soft tissue-windowing can be used for single observer study.

Keywords: atherosclerosis; CCTA; radiomics; repeatability; reproducibility



Citation: Yunus, M.M.; Sabarudin, A.; Karim, M.K.A.; Nohuddin, P.N.E.; Zainal, I.A.; Shamsul, M.S.M.; Yusof, A.K.M. Reproducibility and Repeatability of Coronary Computed Tomography Angiography (CCTA) Image Segmentation in Detecting Atherosclerosis: A Radiomics Study. *Diagnostics* 2022, *12*, 2007. https:// doi.org/10.3390/diagnostics1208 2007

Academic Editors: Albert Comelli, Federica Vernuccio and Roberto Cannella

Received: 3 June 2022 Accepted: 5 August 2022 Published: 19 August 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/).

# 1. Introduction

According to a report from the World Health Organization, in 2019, ischemic heart disease (IHD) was the number one silent killer worldwide, including in Malaysia, with 17.9 million people dying every year due to cardiovascular disease (CVD) [1]. IHD occurs when there is a reduction in the oxygen content carried by the blood vessels to the heart. This can be caused by a blockage by plaque inside the wall area of the blood vessels, known as atherosclerosis. There are three types of atherosclerosis—calcified, semi calcified, and noncalcified plaque [2]. Generally, patients with suspected acute coronary syndrome (ACS) or even without symptoms can undergo the health-screening procedure of coronary computed tomography angiography (CCTA), the invasive coronary angiography (ICA) procedure, and intravascular ultrasound (IVUS) [3]. The radiologist will analyze the image and based on the percentage of blood vessel blockage, classify it as grades 0, 1, 2, 3, or 4 or use the standard reporting system in CCTA, which is CADRADS [4]. CADRADS was developed to standardize the reporting of CCTA, improve communication, and guide therapy. It was published in 2016 by the Society of Cardiovascular Computed Tomography (SCCT), the American College of Radiology (ACR), and the North American Society for Cardiovascular Imaging (NASCI) and has been endorsed by the American College of Cardiology (AC). CCTA has been acknowledged as a noninvasive technique to diagnose and detect atherosclerotic plaques [5,6]. CCTA can highlight the main vessels and identify the plaque morphology, enabling the planning of treatment and estimating the risk stratification in patients with CVD [5–9]. In clinical practice, the radiologist will visualize and interpret the CCTA by assessing the qualitative features of the images. However, this subjective evaluation requires full attention to detail, which may be overlooked if there is too much diagnostic study [10]. As the diagnosis is a subjective assessment highly dependent on the expertise level of the readers, interobserver reproducibility has been found to be low [4]. Hence, precision medicine was established, enabling more precise tools to detect and characterize diseases [11].

In medical imaging, radiomics is a technique that extracts various features from medical images using data algorithms and usually represents them as tumor patterns and other disease characteristics [12,13]. The radiomics process starts with data acquisition, image preprocessing, image segmentation, feature extraction, feature selection, and image classification, as shown in Figure 1. In this research, we were focused more on the first four phases of radiomics study.



Figure 1. Radiomics Process.

The use of radiomics in CCTA radiographs is considered vital as a diagnostic tool, due to its minimal requirements for acquisition and postprocessing [14]. It improves the characterization of plaque-containing blood vessels by using complex and potential algorithms to facilitate the radiologist's decision regarding prospective patient treatment. The use of texture analysis or the so-called part of artificial intelligence (AI) technology is also a promising alternative quantitative analysis to identify and characterize the lesion by using medical imaging information such as from CCTA images [15–19]. Despite its ability, reproducibility in extracting radiomics features remains a concern.

Radiomics has an important role in machine-learning (ML) development and is regarded as the only high-throughput feature for medical images [20]. Radiomics is also subject to the limitation of visual image interpretation, as it covers most of the DICOM format images [21]. Previously, researchers found that data mining from radiographs and predictive analysis widened the scope of medical imaging [22–25]. This can enable prognostic models to classify the pattern and properties of the disease. However, according to Wang et al., the only obstacle to the use of the models is the reliability of their outcome [26].

Challenges during segmentation process of coronary artery in CCTA is inevitable, as the coverage area of the vessel is small and can cause variability in the segmentation result [27]. In addition, conventional manual segmentation is tedious and exhausting, which calls for faster and more reliable segmentation techniques. Previous works evaluated the semiautomatic segmentation methods and found their reliability similar to manual segmentation [28–30]. A work on mammography found that the flood-filling algorithm yields the best result, compared to the watershed and k-means algorithms [31]. Furthermore, the semiautomatic segmentation of the grow-cut algorithm on lung computed tomography (CT) was also found to have high reproducibility and robustness, compared to manual segmentation [32].

Feature extraction is essential to obtain relevant information on input images and represent that information in the lower-dimensionality space [33]. These features are extracted by using an advanced mathematical algorithm that describes phenotypes of lesions that might not be visible to the naked eye. Apart from the selection of segmentation algorithms, medical-image-segmentation accuracy will also be affected by other extrinsic factors such as image analysis, the observers' educational background, image-segmentationrelated experience, and the level of familiarity with the segmentation software [27,34]. To obtain accurate data regarding lesion characteristics, a robust image segmentation algorithm is required. Manual segmentation is widely used in clinical settings, especially for tumour contouring in the oncology sector and volume measurement in the radiology field. However, such segmentation performed through human observation and interpretation was shown to be susceptible to variation, apart from being time-consuming. Hence, this study aimed to evaluate the reproducibility and repeatability of manual and semiautomatic segmentation the atherosclerotic lesions in atherosclerosis of CCTA images. This work aims to identify the most appropriate CCTA image segmentation method for radiomics analysis to quantitatively extract the atherosclerotic lesion.

#### 2. Materials and Methods

The CCTAs with contrast images (axial plane) were retrieved retrospectively from the picture archiving communicating system (PACS) system of Hospital Canselor Tuanku Muhriz (HCTM), Cheras, Kuala Lumpur, Malaysia. The images were collected randomly from December 2019 to March 2021. This observational clinical study was approved by the ethical-committee members on 4 January 2021 (vide approval No. UKM PPI/111/8/JEP-2020-751). Patient consent was waived due to the research involving no more than minimal risk to subjects. Thirty (30) CCTA images of coronary blood vessels (including calcified, noncalcified, and normal blood vessels) were selected using the rule of thumb in the reliability study [35]. The normal coronary artery was identified as a control group. Data were recorded based on random sampling, which is better in reducing bias and more random for optimal data collection results, including radiologist reports and patient demographic data. This research used a single-blinded type of study whereby only the researcher knows the report of the patient. The report was used in the first place to distribute the patient selection based on patient diagnosis. The sample was taken randomly including normal patients, calcified plaque, and noncalcified plaque. Then, two radiology experts were blinded by not knowing the patient's diagnosis and they were both not among those who carried out the reporting on the sample of images. Patient identifications were voided to reduce bias during the image segmentation process. The CCTA images were renamed into sequence of number to avoid patient identification being revealed during the segmentation process. Both manual and semiautomatic segmentation processes were carried out using Local Image Feature Extraction (LIFEx) software (version 7.0.16, Orsay, France) [36]. LIFEx is a free and open-source software platform for analyzing 3-dimensional (3D) medical images, intended for research purposes only, and has not been reviewed or approved by the Food and Drug Administration (FDA) for clinical purposes. Figure 2 shows the overall research workflow in comparing and analyzing the radiomics features that were extracted.



Figure 2. Overall research workflow.

The inclusion criteria included contrast enhanced CCTA with atherosclerosis confirmed by radiologists based on the radiologist report. The exclusion criteria were coronary artery bypass graft (CABG) patients or patients under treatment, stenting, or ballooning, as well as poor-quality CCTA images (with artefacts) and extremely obese patients (body mass index (BMI) more than 35). All the patients were scanned using 320-Multidetector Computed Tomography (CT) Toshiba Aquilion One (Canon Medical Systems, Tochigi, Japan) with a standard prospective electrocardiogram (ECG) protocol. The protocols were pre-set with a tube voltage of 120 kVp, a tube current of 600 mA with automatic mA modulation technique enabled, a gantry rotation time of 350 ms with half gantry turn (half segment reconstruction), a collimation aperture of 0.5 mm  $\times$  160, and 0.537 mm-thickness slices with an overlap distance of 0.5 mm. Ensuring a noise index of 10, the scan size area was 16 cm<sup>2</sup> to allow the entire heart image to be scanned in one gantry rotation in one heartbeat (dynamic volume scan). The scanning range was set between a Z-axis of 120 cm and 160 cm, starting at the bifurcation branch of the trachea until the diaphragm. Throughout the scan, the ECG detection technique was set at 70–80% at the peak of the R-R interval (the interval between one heartbeat and another) [37]. In addition, no table movement per gantry round (pitch) was used because the scan mode used was sequential (step-and-shoot). There was no table movement and overlap of scans during the radiation emitted.

Radiomics feature extractions for manual and semiautomatic segmentation were compared and analyzed. To measure the reproducibility of both (manual and semiautomatic) segmentations, two different methods of viewing bone windows (labeled as: R1; window width (WW): 2500; window level (WL): 480) and soft-tissue windows (labeled as R2-1: window width (WW): 450–1000; window level (WL): 100–300) by two radiology experts were carried out. Then, the manual and semiautomatic segmentations were repeated by the 2nd radiology expert so that we could measure the repeatability using the intraclass correlation (ICC). We labelled it as R2-2 (soft-tissue window setting; window width (WW): 450–1000; window level (WL): 100–300). Both radiologists viewed the coronary blood vessels and segmentized the image relating to the areas of calcified atherosclerosis, noncalcified atherosclerosis, and normal vessels independently. The soft-tissue- and bone-windowing techniques were used by each radiology expert accordingly based on the windowing technique that they commonly used for viewing the CCTA images in their radiology reporting workstation.

Based on the lesion (volume of interest: VOI) areas of the CCTA images on the axial plane, only 3 axial slices of images were chosen to be segmentized. This is to standardize the VOI techniques of segmentation area on each vessel. The LIFEx software extracted the features of atherosclerosis and non-atherosclerosis using a radiomics application into first-order features, shape order features and secondorder features. The interclass and intraclass correlation (ICC) was counted as poor (ICC < 0.5), moderate (0.5 < ICC < 0.75), good (0.75 < ICC < 0.9), and excellent (ICC > 0.9), as referred to in [35]. The details of the research workflow are presented in Figure 3 below.



Figure 3. The framework of the research protocol.

# 2.1. Manual Segmentation (Pencil 2D Technique) Protocol

The loading axial plane of the CCTA data images through the digital imaging and communications in medicine (DICOM) module was loaded into the LIFEx software. Patient identification was removed to reduce any bias-related factors. The observation was conducted under a good room light setting in a private room. Technically, a Windows 10 (64-bit) GPU unit of Intel<sup>®</sup> Core i7 processor with 16GB RAM was used to run these images. Soon

after, the observers identified the location of atherosclerotic plaques and non-atherosclerotic plaques. The DICOM axial images were adjusted with bone window setting (window width (WW): 2500; window level (WL): 480) for radiologist R1. The soft-tissue window setting was set to (WW: 450–1000; WL: 100–300), and was adjusted for radiologist R2-1 and R2-2 (second attempt) independently.

Next, the image was zoomed to 10 times enlargement to focus more on the coronary blood vessel. The pencil 2D technique (manual segmentation) was selected. The nodes were added around the lesion region surrounding the coronary artery using the mouse cursor. The region of interest was then colored in purple (as shown in Figure 4) once the nodes were connected circularly to the first node. This step was repeatedly carried out on two more slices (three slices of axial CCTA in total). The same technique was implemented on two more vessels of coronary arteries.



**Figure 4.** LIFEx software was used to perform manual and semiautomatic segmentation on RCA, LAD, and LCX. (a) Manual segmentation using pencil 2D technique with soft-tissue-windowing setting on RCA (b) The area of segmentation covered after manual segmentation with soft-tissue-windowing setting was performed on RCA.

#### 2.2. Semiautomatic Segmentation (Circle 3D Technique) Protocol

The same technique was repeated using semiautomatic (circle 3D technique also known as growth from seed) segmentation. The chosen nodes were able to adjust the size according to the circular shape of the coronary blood vessel size. The nodes were added around the lesion region surrounding the coronary artery, using the mouse cursor. Subsequently, the flood fill effects were activated, and VOI segmented according to similar voxels' intensity (as shown in Figure 5). To finalize the output, the segmented lesion was manually edited either to delete some extra covered area or to add an area not covered by the semiautomatic cursor in finalizing the semiautomatic segmentation.

The flood-filled algorithm is capable of connecting the area in the multidimensional array by allowing the related intensity voxels to access the selected node, which is decided by the users. This algorithm is comparable with the bucket tool in paint programs, which fills connected, similar-intensity voxels with a different color [31]. The algorithm was initiated with the first node during the selection of the volume of interest (VOI). The pixels were connected directionally from the first node to the former. Subsequently, the intensity voxels were decided, and the algorithm detected the identified path and switched to a different color. To reduce the leakage effect, the VOI was manipulated using the neighborhood size parameter.



**Figure 5.** Semiautomatic segmentation using circle 3D technique in LIFEx software in bone window setting on the coronary blood vessel.

#### 2.3. Feature Extraction

Once the segmentation process was finished, the feature was extracted using the textural extraction analysis in the LIFEx software, with a bin size of 10. All features were extracted based on the mathematical algorithm predicated on pixel intensities. The first-order, shape order, and second (2nd)-order features were extracted to a Microsoft Excel worksheet (CSV. formatted file). Of each of the 30 images, we repeated 6 times the image segmentation where each technique produced three types of radiomics features which were shape, texture, and lesion intensity. We managed to extract 11,700 radiomic features of the CCTA images to allow quantification of the atherosclerotic and non-atherosclerotic lesion characteristics. Table 1 summarizes the composition of features which are divided into three groups: (I) intensity (1st-order features), (II) shape order features, and (III) textural (2nd-order features). The total features extracted from the volume of interest for lesion intensity, shape, and texture were 29, 5, and 31, respectively.

Features $(n = 180)$	<b>Radiomics Features</b>
Lesion intensity (1st-order features)	29  imes 180
Shape order features	5  imes 180
Texture (2nd-order features)	31  imes 180

The first-order statistics could distinguish the histogram of voxel intensity within the atherosclerotic lesion region on CCTA images. Furthermore, the shape features were able to describe the volume properties of the lesion. The patterns or spatial distributions of voxel intensities derived from the gray-level dependence matrix (GLDM), gray-level co-occurrence matrix (GLCM), and gray-level run-length matrices (GLRLM) were defined by the textural features. The features from the co-occurrence and run-length matrices were estimated by averaging all the 13 symmetric directions in three dimensions [38].

Several interest features, such as entropy, contrast, uniformity, and correlation, are presented in the equation below:

$$Entropy = -\sum_{i=1}^{N_g} p(i) \log_2(p(i) + \epsilon)$$
(1)

Entropy measures the average amount of information required to encode the image values, where  $N_g$  = the number of non-zero bins and p(i) = normalized first-order histogram.

Contrast = 
$$\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i-j)^2 p(i,j)$$
 (2)

Contrast determines the local intensity variation present in the image. A larger value correlates with the greater disparity in intensity values among neighboring voxels.

Uniformity = 
$$\sum_{i=1}^{N_g} p(i)^2$$
 (3)

Uniformity measures the sum of squares of each intensity value. A greater uniformity implies greater homogeneity. Correlation defines the linear dependency of gray-level values to their respective voxels in GLCM. The value for correlation is between 0 (uncorrelated) and 1 (perfectly correlated). The equation of correlation is shown below:

$$Correlation = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j)ij - \mu_x \mu_y}{\sigma_x(i)\sigma_v(j)}$$
(4)

where  $\mu_x$ ,  $\mu_y$  and  $\sigma_x$ ,  $\sigma_y$  are mean gray-level intensity and standard deviation of  $p_x$  and  $p_y$ , respectively. Table 2 summarizes the radiomic characteristics derived in this investigation in detail.

### 2.4. Statistical Analysis

The intraclass correlation coefficient (intra-CC) represents the correlations within the class of data and is used to determine the repeatability of the extracted features, while the reproducibility, which was determined by the interclass correlation (inter-CC) analysis, was used in this study. Briefly, the relationship between the observer was measured and categorized into three models which can be chosen appropriately, depending on the experimental situation. In this study, variance values were estimated to define the interobserver segmentations by using a two-way mixed-effect model of analysis of variance (ANOVA) [24]. Mathematically, ICC is defined as:

ICC (A, 1) = 
$$\frac{MS_{R} - MS_{E}}{MS_{R} + (k+1)MS_{g} + \frac{k}{n}(MS_{C} - MS_{E})}$$
(5)

One-way analysis of variance (ANOVA) was used to obtain the ICC values for intra observer segmentation [14,15]. The equation below defines ICC (C,1):

$$ICC(C,1) = \frac{MS_R - MS_W}{MS_R + (k-1)MS_W}$$
(6)

where  $MS_R$  = mean square for rows,  $MS_W$  = mean square for residual sources of variance,  $MS_E$  = mean square error,  $MS_C$  = mean square for columns, and k and n are the numbers of observers involved and subjects.

The reproducibility evaluation was performed by letting one observer segmentize 30 images 2 weeks apart. The measurement was based on multiple initializations of the segmentation algorithm from the same observer. Interobserver reproducibility was determined from several observers according to the segmentation technique and the degree of agreement between observers. The Wilcoxon rank-sum test with a *p*-value set at 0.05 was used to define the differences in reproducibility. We used the Statistical Package for Social Sciences (SPSS, also known as IBM SPSS statistics) version 25 (SPSS, Chicago, IL, USA) to analyze the data and all data were expressed in (mean  $\pm$  standard deviation).

**Table 2.** List of first-order features (n = 29), second-order features (n = 31), and shape features (n = 5). Reprinted/adapted with permission from Ref. [15].

# 3. Results

# 3.1. Descriptive Analysis

Of 30 images of CCTA, 50% were categorized as coronary artery disease—reporting and data system (CADRADS)-4, 30% as CADRADS 2, 10% as CADRADS 1, and 10% as CADRADS 5. In total, 36.7% were calcified atherosclerosis, 23.3% were noncalcified plaque, and 40% were normal.

# 3.2. ICC for Manual and Semiautomatic Segmentation for First (1st)-Order and Shape Order Features

Both inter- and intra class coefficients (ICC) for manual and semi-automatic segmentation of the first-order and shape order radiomics features were compared by the means and standard deviation, as shown in Figures 6 and 7. The first 29 radiomics features were categorized as the first-order features followed by five types of shape order features. These radiomics features were then compared to their ICC to find their reproducibility and repeatability on both types of segmentation which were manual and semiautomatic. As we can visually analyze from the pattern of the bar chart graphs, we can see that the majority (52.9% (18/34)) of semiautomatic segmentation was higher in reproducibility as compared to manual segmentation, while the majority of the first-order and shape order features showed higher repeatability on manual segmentation (82.4% (28/34)).



**Figure 6.** Interclass Correlation Coefficients (inter-CC) between manual and semi-automatic segmentation for 1st order features including the shape features to compare its reproducibility.

### 3.3. ICC for Manual and Semiautomatic Segmentation for Second-Order Features

The ICC for manual and semiautomatic segmentation of the second-order radiomics features was compared by the means and standard deviation as shown in Figures 8 and 9. There were 31 radiomics features of the second order which were extracted from the CCTA images. These radiomics features were then compared to their ICC to find their reproducibility and repeatability on both types of segmentation which were manual and semiautomatic. The bar chart graphs show the majority (71% (22/31)) of semiautomatic segmentation was higher in reproducibility as compared to manual segmentation, while the majority of the second-order radiomic features showed higher repeatability on manual segmentation (83.9% (26/31)).



**Figure 7.** Intraclass Correlation Coefficient (intra-CC) between manual and semi-automatic segmentation for 1st order Features including the Shape Features to compare its repeatability.



**Figure 8.** Interclass Correlation Coefficient (inter-CC) between manual and semi-automatic segmentation for 2nd order features to compare its reproducibility.

Figure 10 shows the overall result in a boxplot graph of comparing the reproducibility and repeatability between manual and semiautomatic segmentation for all radiomics features. It is clearly shown that the highest repeatability was in manual segmentation for the first-order and shape order features with an excellent median of ICC = 0.961 (p < 0.001), and ICC = 0.924 (p < 0.001) for the second-order features. The average mean of repeatability and reproducibility for the first-order and shape features' (Cronbach's alpha) measured ICC was 0.979 (p < 0.001) and 0.859 (p < 0.001) for the second-order features. Meanwhile, the highest reproducibility in semiautomatic segmentation for the first-order and shape order features had an excellent median of ICC 0.920 (p < 0.001).



**Figure 9.** Intraclass Correlation Coefficient (inter-CC) between manual and semi-automatic segmentation for 2nd order features to compare its repeatability.



**Figure 10.** Repeatability and reproducibility of manual and semi-automatic segmentation for the first-, shape and second-order features. The (\*) shows the highest repeatability and reproducibility mean score with a significant *p*-value (p < 0.001).

#### 3.4. ICC Level on Reproducibility and Repeatability for Manual and Semiautomatic Segmentation

We categorized both ICCs as poor (ICC < 0.5), moderate (0.5 < ICC < 0.75), good (0.75 < ICC < 0.9), or excellent (ICC > 0.9) [35]. Table 3 shows the reproducibility and repeatability of manual and semi-automatic segmentation for the first-order, shape order, and second-order features. Consequently, semiautomatic segmentation had the highest percentage of excellent (ICC > 0.9) reproducibility for the segmentation of coronary blood vessels with atherosclerotic lesions and normal blood vessels for first-order features. Meanwhile, manual segmentation had the highest percentage of excellent (ICC > 0.9) reproducibility for segmentation also gained the highest percentage of (ICC > 0.9) repeatability for segmentation with 74% for first-order features and 68% (excellent ICC > 0.9) for second-order features. Comparably, the manual and semiautomatic segmentations both yielded a majority of excellent (ICC > 0.9) repeatability and reproducibility based on the first-order, shape order, and second-order features.

**Table 3.** Comparison of reproducibility and repeatability of Manual and Semiautomatic Segmentation based on ICC levels (\* shows the highest % for excellent ICC).

Radiomics Features	ICC Level	Type of ICC	Manual	Semi- Automatic
First-order and shape order	Excellent (ICC > 0.9)	Reproducibility (inter-CC)	19 (56%)	20 (59%) *
		Repeatability (intra-CC)	25 (74%) *	18 (53%)
	Good (0.75 < ICC < 0.9)	Reproducibility (inter-CC)	4 (12%)	7 (21%)
		Repeatability (intra-CC)	5 (15%)	5 (15%)
	Moderate (0.5 < ICC < 0.75)	Reproducibility (inter-CC)	5 (15%)	4 (12%)
		Repeatability (intra-CC)	4 (12%)	9 (26%)
	Low (ICC < 0.5)	Reproducibility (inter-CC)	6 (18%)	3 (9%)
		Repeatability (intra-CC)	0 (0%)	2 (6%)
Second order	Excellent (ICC > 0.9)	Reproducibility (inter-CC)	11 (35%) *	10 (32%)
		Repeatability (intra-CC)	21 (68%) *	13 (42%)
	Good (0.75 < ICC < 0.9)	Reproducibility (inter-CC)	10 (32%)	9 (29%)
		Repeatability (intra-CC)	10 (32%)	12 (39%)
	Moderate (0.5 < ICC < 0.75)	Reproducibility (inter-CC)	6 (19%)	8 (26%)
		Repeatability (intra-CC)	0 (0%)	6 (19%)
	Low (ICC < 0.5)	Reproducibility (inter-CC)	4 (13%)	4 (13%)
		Repeatability (intra-CC)	0 (0%)	0 (0%)

#### 4. Discussion

This study shows that manual segmentation yielded a high repeatability result since the same observer was using the same method and windowing technique during the image segmentation process. This result shows that using a soft-tissue-windowing setting in viewing the CCTA images for image segmentation purposes was able to produce high repeatability. Yet, there are still some disadvantages to the manual technique which is monotonous, time-consuming, and did affect the interobserver variability [39] as shown in the result above.

Meanwhile, semi-automatic segmentation is proven to have higher reproducibility in CCTA images segmentation, using two different windowing techniques. Other than higher reproducibility result, semi-automatic segmentation also taking less time in segmentation as compared to the manual segmentation [39]. This result also shows that the windowing technique also gave impact on reproducing the same segmentation result in manual segmentation. It is suggested to maintain a standard image viewing setting as it may affect

the image visualization, especially in detecting atherosclerosis. This result is supported by the previous study conducted on other imaging modalities, diseases, and techniques of segmentation [40].

Although semiautomated segmentation was proved less susceptible to intra- and interobserver variability compared to manual delineation [41,42], it may still cause contrasting results even using the same segmentation approach. This may be due to the requirement of human interaction in semiautomated segmentation; different initialization points' locations and sizes may cause different outcomes [43]. This was also supported by Owen in 2013 [44] who found that with less human interaction during the segmentation process, a more reproducible result could be obtained. Other than that, they suggested using the same segmentation software and approach in obtaining radiomics information because significant inter-software variability was observed in their study. The same suggestion was also mentioned in theother study [45].

The advantages of semiautomated segmentation include allowing the viewer to detect those areas of interest which have about the same pixel value beside it. This has been proven beneficial to apply, especially in detecting disease in small areas [34]. This will be one of the reasons why semi-automated segmentation has lower repeatability as compared to manual segmentation.

Notably, the majority of the first-order, shape order and second-order features for both manual and semiautomatic segmentation had excellent reproducibility and repeatability with (ICC > 0.9). Yet, the average percentage of excellent level of reproducibility and repeatability (ICC > 0.9) was higher in the first-order and shape order features. This is due to the technical approach of measuring the first-order and second-order features itself, where the first-order features only focus on the mathematical calculation of the Hounsfield unit in one single pixel, while the second-order features are prone to measure the pixel next to another pixel of the image to be measured. Yet, these radiomics features will be highly beneficial in supervising various automated machine-learning models, especially in detecting any diseases in medical images particularly at the small area. This result is supported by the previous study [15,16,46].

The segmentation process is very crucial in radiomics study. If the area of segmentation is carried out wrongly, it may affect the machine-learning codes to classify diseases from the CCTA images. Therefore, this research is very prominent to enhance the importance of choosing the best segmentation technique in comparing with two different windowing techniques in CT scan. The data that we managed to extract out from the CCTA images were consisted of large number of patients' diagnostic informations which will be very helpful in predicting the disease quantitatively by usingproper image windowing and segmentationtechnique.

Meanwhile, this research simulated the segmentation techniques in two different centers in Malaysia. With reference to this study, it is hoped that future research can apply the semi-automated segmentation in soft-tissue windowing setting accordingly. A previous study also was conducted on CT lung images as a reference [47]. This may help radiologists to refer to the standard guideline in viewing the CCTA images, especially in image segmentation of coronary blood vessels. This will give a large contribution to patient management, hospital management, and finally, to the nation in detecting atherosclerosis. It is suggested for future research to further analyze the radiomics features of two different segmentation techniques by measuring the accuracy of image classification in detecting atherosclerosis using a machine-learning algorithm. Therefore, it may give a higher impact in terms of the accuracy in detecting atherosclerosis before implementation to a real prospective patient.

The limitation of this study is that it involved only two radiology experts from two different centers. It can also be improved in the next reliability research study by adding a minimum of three radiologists to interpret the images [35]. Yet, this research is highly impactful in studying the method and protocol that can be implemented in a future study.

# 5. Conclusions

This study demonstrated the manual radiomics feature extraction with soft-tissuewindowing setting yielded higher repeatability as compared to semiautomatic segmentation, with 82.4% for first-order/shape Features and 83.9% for second-order features. Meanwhile, semiautomatic segmentation had higher reproducibility as compared to manual segmentation, with 71% for first-order/shape features and 52.9% for second-order features. The majority of the first-order and second-order features for both manual and semiautomatic segmentation had an excellent repeatability and reproducibility (ICC > 0.9). In conclusion, semiautomated segmentation is recommended for interobserver study while manual segmentation with soft-tissue windowing can be used for single-observer study.

Author Contributions: Formal analysis, M.M.Y. and M.K.A.K.; investigation, A.S. and M.K.A.K.; methodology, M.S.M.S.; project administration, A.K.M.Y.; resources, I.A.Z. and M.S.M.S.; supervision, A.S.; validation, P.N.E.N. and A.K.M.Y.; visualization, I.A.Z.; writing—original draft, M.M.Y.; writing—review and editing, M.K.A.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by Ministry of Higher Education and Universiti Putra Malaysia with grant number FRGS/1/2020/STG07/UPM/02/3.

**Institutional Review Board Statement:** The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board (or Ethics Committee) of Hospital Canselor Tuanku Muhriz (HCTM), (UKM PPI/111/8/JEP-2020-751; 4 January 2021).

**Informed Consent Statement:** Patient consent was waived due to the research involving no more than minimal risk to subjects.

Data Availability Statement: Not applicable.

Acknowledgments: Hospital Canselor Tuanku Muhriz (HCTM), Institut Jantung Negara (IJN), Universiti Kebangsaan Malaysia (UKM), and Universiti Putra Malaysia (UPM).

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

# References

- WHO. The Top 10 Causes of Death. Available online: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causesof-death (accessed on 2 June 2022).
- Kolossváry, M.; Szilveszter, B.; Merkely, B.; Maurovich-Horvat, P. Plaque Imaging with CT-A Comprehensive Review on Coronary CT Angiography Based Risk Assessment. *Cardiovasc. Diagn. Ther.* 2017, 7, 489–506. [CrossRef]
- Munnur, R.K.; Cameron, J.D.; Ko, B.S.; Meredith, I.T.; Wong, D.T.L. Cardiac CT: Atherosclerosis to Acute Coronary Syndrome. *Cardiovasc. Diagn. Ther.* 2014, 4, 430–448. [CrossRef] [PubMed]
- Cury, R.C.; Abbara, S.; Achenbach, S.; Agatston, A.; Berman, D.S.; Budoff, M.J.; Dill, K.E.; Jacobs, J.E.; Maroules, C.D.; Rubin, G.D.; et al. CAD-RADSTM Coronary Artery Disease—Reporting and Data System. An Expert Consensus Document of the Society of Cardiovascular Computed Tomography (SCCT), the American College of Radiology (ACR) and the North American Society for Cardiovascular Imaging (NASCI). Endorsed by the American College of Cardiology. J. Cardiovasc. Comput. Tomogr. 2016, 10, 269–281. [CrossRef] [PubMed]
- Newby, D.; Williams, M.; Hunter, A.; Pawade, T.; Shah, A.; Flapan, A.; Forbes, J.; Hargreaves, A.; Leslie, S.; Lewis, S.; et al. CT Coronary Angiography in Patients with Suspected Angina Due to Coronary Heart Disease (SCOT-HEART): An Open-Label, Parallel-Group, Multicentre Trial. *Lancet* 2015, 385, 2383–2391. [CrossRef]
- 6. Cury, R.C.; Budoff, M.; Taylor, A.J. Coronary CT Angiography versus Standard of Care for Assessment of Chest Pain in the Emergency Department. J. Cardiovasc. Comput. Tomogr. 2013, 7, 79–82. [CrossRef] [PubMed]
- Douglas, C.; Hoffmann, U.; Patel, M.R.; Mark, D.B.; Al-Khalidi, H.R.; Cavanaugh, B.; Cole, J.; Douglas, P.S.; Dolor, R.J.; Fordyce, C.B.; et al. Outcomes of Anatomical Versus Functional Testing for Coronary Artery Disease. *N. Engl. J. Med.* 2015, 372, 1291–1300. [CrossRef]
- Karim, M.K.A.; Sabarudin, A.; Muhammad, N.A.; Ng, K.H. A Comparative Study of Radiation Doses between Phantom and Patients via CT Angiography of the Intra-/Extra-Cranial, Pulmonary, and Abdominal/Pelvic Arteries. *Radiol. Phys. Technol.* 2019, 12, 1–8. [CrossRef]
- Harun, H.H.; Abdul Karim, M.K.; Abd Rahman, M.A.; Abdul Razak, H.R.; Che Isa, I.N.; Harun, F. Establishment of CTPA Local Diagnostic Reference Levels with Noise Magnitude as a Quality Indicator in a Tertiary Care Hospital. *Diagnostics* 2020, 10, 680. [CrossRef]

- Haniff, N.S.M.; Karim, M.K.B.A.; Ali, N.S.; Rahman, M.A.A.; Osman, N.H.; Saripan, M.I. Magnetic Resonance Imaging Radiomics Analysis for Predicting Hepatocellular Carcinoma. In Proceedings of the 2021 International Congress of Advanced Technology and Engineering (ICOTEN), Taiz, Yemen, 4–5 July 2021; pp. 1–5. [CrossRef]
- 11. Kolossváry, M.; De Cecco, C.N.; Feuchtner, G.; Maurovich-Horvat, P. Advanced Atherosclerosis Imaging by CT: Radiomics, Machine Learning and Deep Learning. *J. Cardiovasc. Comput. Tomogr.* **2019**, *13*, 274–280. [CrossRef]
- Lambin, P.; Rios-Velazquez, E.; Leijenaar, R.; Carvalho, S.; Van Stiphout, R.G.P.M.; Granton, P.; Zegers, C.M.L.; Gillies, R.; Boellard, R.; Dekker, A.; et al. Radiomics: Extracting More Information from Medical Images Using Advanced Feature Analysis. *Eur. J. Cancer* 2012, *48*, 441–446. [CrossRef]
- Rizzo, S.; Botta, F.; Raimondi, S.; Origgi, D.; Fanciullo, C.; Morganti, A.G.; Bellomi, M. Radiomics: The Facts and the Challenges of Image Analysis. *Eur. Radiol. Exp.* 2018, 2, 36. [CrossRef] [PubMed]
- 14. Xu, P.; Xue, Y.; Schoepf, U.J.; Varga-Szemes, A.; Griffith, J.; Yacoub, B.; Zhou, F.; Zhou, C.; Yang, Y.; Xing, W.; et al. Radiomics: The Next Frontier of Cardiac Computed Tomography. *Circ. Cardiovasc. Imaging* **2021**, *14*, 256–264. [CrossRef] [PubMed]
- Yunus, M.M.; Khairuddin, A.; Yusof, M.; Zaidi, M.; Rahman, A.; Koh, X.J.; Sabarudin, A.; Nohuddin, P.N.E.; Ng, K.H.; Mustafa, M.; et al. Automated Classification of Atherosclerotic Radiomics Features in Coronary Computed Tomography Angiography (CCTA). *Diagnostics* 2022, 12, 1660. [CrossRef]
- Yunus, M.M.; Sabarudin, A.; Hamid, N.I.; Yusof, A.K.M.; Nohuddin, P.N.E.; Karim, M.K.A. Automated Classification of Atherosclerosis in Coronary Computed Tomography Angiography Images Based on Radiomics Study Using Automatic Machine Learning. In Proceedings of the 2022 International Conference on Electronics and Renewable Systems (ICEARS), Tuticorin, India, 16–18 March 2022; pp. 1895–1903. [CrossRef]
- Kolossváry, M.; Karády, J.; Kikuchi, Y.; Ivanov, A.; Schlett, C.L.; Lu, M.T.; Foldyna, B.; Merkely, B.; Aerts, H.J.; Hoffmann, U.; et al. Radiomics versus Visual and Histogram-Based Assessment to Identify Atheromatous Lesions at Coronary CT Angiography: An Ex Vivo Study. *Radiology* 2019, 293, 89–96. [CrossRef] [PubMed]
- Rajendra Acharya, U.; Meiburger, K.M.; Wei Koh, J.E.; Vicnesh, J.; Ciaccio, E.J.; Shu Lih, O.; Tan, S.K.; Aman, R.R.A.R.; Molinari, F.; Ng, K.H. Automated Plaque Classification Using Computed Tomography Angiography and Gabor Transformations. *Artif. Intell. Med.* 2019, 100, 101724. [CrossRef]
- Candemir, S.; White, R.D.; Demirer, M.; Gupta, V.; Bigelow, M.T.; Prevedello, L.M.; Erdal, B.S. Automated Coronary Artery Atherosclerosis Detection and Weakly Supervised Localization on Coronary CT Angiography with a Deep 3-Dimensional Convolutional Neural Network. *Comput. Med. Imaging Graph.* 2020, *83*, 101721. [CrossRef]
- Çinarer, G.; Gürsel, B.; Haşim, A. Prediction of Glioma Grades Using Deep Learning with Wavelet Radiomic Features. *Appl. Sci.* 2020, 10, 6296. [CrossRef]
- Fairuz, S.; Radzi, M.; Khalis, M.; Karim, A.; Saripan, M.I.; Amiruddin, M.; Rahman, A.; Nurzawani, I.; Isa, C.; Ibahim, M.J.; et al. Hyperparameter Tuning and Pipeline Optimization via Grid Search Method and Tree-Based AutoML in Breast Cancer Prediction. J. Pers. Med. 2021, 11, 978. [CrossRef]
- Cao, X.H.; Stojkovic, I.; Obradovic, Z. A Robust Data Scaling Algorithm to Improve Classification Accuracies in Biomedical Data. BMC Bioinform. 2016, 17, 1–10. [CrossRef]
- El-Dahshan, E.A.S.; Mohsen, H.M.; Revett, K.; Salem, A.B.M. Computer-Aided Diagnosis of Human Brain Tumor through MRI: A Survey and a New Algorithm. *Expert Syst. Appl.* 2014, 41, 5526–5545. [CrossRef]
- Zhou, T.; Ruan, S.; Canu, S. A Review: Deep Learning for Medical Image Segmentation Using Multi-Modality Fusion. Array 2019, 3–4, 100004. [CrossRef]
- Simi, V.R.; Joseph, J. Segmentation of Glioblastoma Multiforme from MR Images—A Comprehensive Review. Egypt. J. Radiol. Nucl. Med. 2015, 46, 1105–1110. [CrossRef]
- Wang, L.; Tan, J.; Ge, Y.; Tao, X.; Cui, Z.; Fei, Z.; Lu, J.; Zhang, H.; Pan, Z. Assessment of Liver Metastases Radiomic Feature Reproducibility with Deep-Learning-Based Semi-Automatic Segmentation Software. *Acta Radiol.* 2021, 62, 291–301. [CrossRef] [PubMed]
- Kolossváry, M.; Jávorszky, N.; Karády, J.; Vecsey-Nagy, M.; Dávid, T.Z.; Simon, J.; Szilveszter, B.; Merkely, B.; Maurovich-Horvat, P. Effect of Vessel Wall Segmentation on Volumetric and Radiomic Parameters of Coronary Plaques with Adverse Characteristics. J. Cardiovasc. Comput. Tomogr. 2020, 15, 137–145. [CrossRef] [PubMed]
- McGraw, K.O.; Wong, S.P. "Forming Inferences about Some Intraclass Correlations Coefficients": Correction. *Psychol. Methods* 1996, 1, 390. [CrossRef]
- 29. Zhao, S.; Ren, W.; Zhuang, Y.; Wang, Z. The Influence of Different Segmentation Methods on the Extraction of Imaging Histological Features of Hepatocellular Carcinoma CT. J. Med. Syst. 2019, 43, 101. [CrossRef]
- Mat Radzi, S.F.; Abdul Karim, M.K.; Saripan, M.I.; Abd Rahman, M.A.; Osman, N.H.; Dalah, E.Z.; Mohd Noor, N. Impact of Image Contrast Enhancement on Stability of Radiomics Feature Quantification on a 2D Mammogram Radiograph. *IEEE Access* 2020, *8*, 127720–127731. [CrossRef]
- Podgornova, Y.A.; Sadykov, S.S. Comparative Analysis of Segmentation Algorithms for the Allocation of Microcalcifications on Mammograms. In Proceedings of the CEUR Workshop Proceedings, Pescaia, Italy, 16–19 June 2019; Volume 2391, pp. 122–127.
- Qiu, Q.; Duan, J.; Gong, G.; Lu, Y.; Li, D.; Lu, J.; Yin, Y. Reproducibility of Radiomic Features with GrowCut and GraphCut Semiautomatic Tumor Segmentation in Hepatocellular Carcinoma. *Transl. Cancer Res.* 2017, 6, 940–948. [CrossRef]

- Kumar, G.; Bhatia, P.K. A Detailed Review of Feature Extraction in Image Processing Systems. In Proceedings of the 2014 Fourth International Conference on Advanced Computing & Communication Technologies, Rohtak, India, 8–9 February 2014; pp. 5–12. [CrossRef]
- Haniff, N.S.M.; Abdul Karim, M.K.; Osman, N.H.; Saripan, M.I.; Che Isa, I.N.; Ibahim, M.J. Stability and Reproducibility of Radiomic Features Based Various Segmentation Technique on MR Images of Hepatocellular Carcinoma (HCC). *Diagnostics* 2021, 11, 1573. [CrossRef]
- Koo, T.K.; Li, M.Y. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J. Chiropr. Med. 2016, 15, 155–163. [CrossRef]
- Nioche, C.; Orlhac, F.; Boughdad, S.; Reuze, S.; Goya-Outi, J.; Robert, C.; Pellot-Barakat, C.; Soussan, M.; Frouin, F.; Buvat, I. LIFEx: A Freeware for Radiomic Feature Calculation in Multimodality Imaging to Accelerate Advances in the Characterization of Tumor Heterogeneity. *Cancer Res.* 2018, 78, 4786–4789. [CrossRef]
- Sabarudin, A.; Siong, T.W.; Chin, A.W.; Hoong, N.K.; Karim, M.K.A. A Comparison Study of Radiation Effective Dose in ECG-Gated Coronary CT Angiography and Calcium Scoring Examinations Performed with a Dual-Source CT Scanner. *Sci. Rep.* 2019, *9*, 4374. [CrossRef] [PubMed]
- Xu, L.; Gao, Q.; Yousefi, N. Brain Tumor Diagnosis Based on Discrete Wavelet Transform, Gray-Level Co-Occurrence Matrix, and Optimal Deep Belief Network. *Simulation* 2020, *96*, 867–879. [CrossRef]
- Dias, R.D.; Shah, J.A.; Zenati, M.A. Artificial Intelligence in Cardiothoracic Surgery. *Minerva Cardioangiol.* 2020, 68, 532–538.
   [CrossRef]
- 40. Das, M.J.; Mahanta, L.B. Lung Segmentation from CT Images: Impact of Different Window Settings on the Accuracy of Segmentation. J. Emerg. Technol. Innov. Res. 2019, 5, 189–195.
- Bendtsen, C.; Kietzmann, M.; Korn, R.; Mozley, P.D.; Schmidt, G.; Binnig, G. X-ray Computed Tomography: Semiautomated Volumetric Analysis of Late-Stage Lung Tumors as a Basis for Response Assessments. *Int. J. Biomed. Imaging* 2011, 2011, 11. [CrossRef]
- 42. Velazquez, E.R.; Parmar, C.; Jermoumi, M.; Mak, R.H.; Van Baardwijk, A.; Fennessy, F.M.; Lewis, J.H.; De Ruysscher, D.; Kikinis, R.; Lambin, P.; et al. Volumetric CT-Based Segmentation of NSCLC Using 3D-Slicer. *Sci. Rep.* **2013**, *3*, 3529. [CrossRef]
- Chen, B.; Zhang, R.; Gan, Y.; Yang, L.; Li, W. Development and Clinical Application of Radiomics in Lung Cancer. *Radiat. Oncol.* 2017, 12, 1–8. [CrossRef]
- 44. Owens, C.A.; Peterson, C.B.; Tang, C.; Koay, E.J.; Yu, W.; Mackin, D.S.; Li, J.; Salehpour, M.R.; Fuentes, D.T.; Court, L.E.; et al. Lung Tumor Segmentation Methods: Impact on the Uncertainty of Radiomics Features for Non-Small Cell Lung Cancer. *PLoS ONE* **2018**, *13*, e0205003. [CrossRef]
- 45. Kalpathy-Cramer, J.; Zhao, B.; Goldgof, D.; Gu, Y.; Wang, X.; Yang, H.; Tan, Y.; Gillies, R.; Napel, S. A Comparison of Lung Nodule Segmentation Algorithms: Methods and Results from a Multi-Institutional Study. J. Digit. Imaging 2016, 29, 476–487. [CrossRef]
- Kim, G.Y.; Lee, J.H.; Hwang, Y.N.; Kim, S.M. A Novel Intensity-Based Multi-Level Classification Approach for Coronary Plaque Characterization in Intravascular Ultrasound Images. *BioMed. Eng. Online* 2018, 17, 151. [CrossRef] [PubMed]
- Mohd Yunus, M.; Hui Sin, N.; Sabarudin, A.; Abdul Karim, M.K.; Ar, R.; Mohd Shamsul, M.S. Comparative Study of the Manual and Semi- Automated Segmentation Technique in Computed Tomography (CT) Lung Cancer: A Radiomics Study. *J. Med. Heal.* 2021, 16, 1–62.