Expanding the deep-learning model to diagnosis LVNC: limitations and trade-offs

G. Bernabé, P. González-Férez, J. M. García, G. Casas & J. González-Carrillo

Published online: 11 Feb 2024.

Submit your article to this journal ⬀

View related articles ⬀

View Crossmark data ⬀

Taylor & Francis
Taylor & Francis Group

🔓 OPEN ACCESS | Check for updates

Expanding the deep-learning model to diagnosis LVNC: limitations and trade-offs

G. Bernabé [ID][a], P. González-Férez [ID][a], J. M. García [ID][a], G. Casas [ID][b] and J. González-Carrillo [ID][c]

[a]Computer Engineering Department, University of Murcia, Murcia, Spain; [b]Cardiology, Hospital Universitari Vall d'Hbron, Barcelona, Spain; [c]Cardiology, Hospital Virgen de la Arrixaca, Murcia, Spain

ABSTRACT

Hyper-trabeculation or non-compaction in the left ventricle of the myocardium (LVNC) is a recently classified form of cardiomyopathy. Several methods have been proposed to quantify the trabeculae accurately in the left ventricle, but there is no general agreement in the medical community to use a particular approach. In the previous work, we proposed DL-LVTQ, a deep-learning approach for left ventricular trabecular quantification based on a U-Net CNN architecture. In this work, we have extended and adapted DL-LVTQ to cope with patients with different particularities and cardiomyopathies. Patient images were taken from different scanners and hospitals. We have modified and adapted the U-Net convolutional neural network to account for the different particularities of a heterogeneous group of patients with multiple cardiomyopathies and inherited cardiomyopathies. The inclusion of new groups of patients has increased the accuracy, specificity and Kappa values while maintaining the sensitivity of the proposed method. Therefore, a better-prepared diagnosis tool is ready for various cardiomyopathies with different characteristics. Cardiologists have considered that 98.9% of the evaluated outputs are verified clinically for diagnosis. Therefore, the high precision to segment the different cardiac structures allows us to make a robust diagnostic system objective and faster, decreasing human error and time spent.

1. Introduction

According to the World Health Organization reports (World Health Organization2022 2022; Mc Namara et al. 2019), cardiovascular diseases are one of the leading causes of death globally, causing about 32% of all deaths worldwide. Among cardiovascular diseases, Left Ventricular Non-Compaction (LVNC) is a recently classified form of cardiomyopathy characterised by abnormal trabeculations or non-compacted tissue in the left ventricle cavity (Towbin et al. 2015), that can be found in association with other cardiomyopathies (Udeoji et al. 2013; Arbustini et al. 2014; Biagini et al. 2006).

Several methods based on magnetic resonance imaging (MRI) have been proposed to accurately quantify the trabeculae in the left ventricle (LV) of the myocardium (Jacquier et al. 2010; Captur et al. 2013, 2014; Choi et al. 2016; Bernabé et al. 2017). However, there is a disagreement in the cardiology community to determine a universal standard accompanied by excessive time to obtain the surfaces manually and the subjectivity of the cardiologist to carry it out. An understandable measure to diagnose this cardiomyopathy is to calculate the percentage of the trabecular volume to the total volume of the non-compacted wall of the left ventricle ($VT\%$).

In the last decade, Deep Neural Networks (DNNs) have become extremely popular, and one of the main methods of modern Artificial Intelligence (AI) (Sze et al. 2017). DNNs are widely used in many domains, and the number of applications that use them has significantly increased. DNNs are employed in many scientific fields, such as image recognition, speech recognition, or autonomous vehicles, and other medical areas, such as cancer detection or protein folding prediction. In most of these domains, DNNs can outperform human capabilities thanks to their ability to perform high-level abstractions from large datasets (Sze et al. 2017).

In fact, several works have recently proposed an automatic solution based on deep learning (DL) techniques to determine the left ventricle volume through MRI (Litjens et al. 2017; Chen et al. 2020), a crucial issue in assessing cardiac diseases. In addition, by using high-performance computing, recent publications exploit various possibilities of deep learning to segment the left and the right ventricle (Pérez-Pelegrí et al. 2021; Penso et al. 2021; Li et al. 2021). Moreover, Bartoli et al. (2020) have proposed a deep learning framework to estimate LVNC based on a DenseNet convolutional neural network (CNN) architecture; however, the trabeculae area is not assessed with enough precision.

Our research group has worked on this topic intensively. In previous works, we proposed a semi-automatic software (QLVTHC) (Bernabé et al. 2020, 2021) and an automatic tool (SOST) ([Bernabé et al. 2018]) that delineate segmentations of the endocardium border and trabecular masses on cardiovascular magnetic resonances (CMR). Both proposals were based on traditional computer vision techniques. Lastly, we have proposed an automatic tool called Deep Learning for Left Ventricular Trabecular Quantification (DL-LVTQ) based on the usage of Convolutional Neural Networks (CNNs) (Rodríguez-de-Vera et al. 2022). In particular, this proposal uses as CNN model the well-known U-Net architecture (Ronneberger et al. 2015) that provides fast and precise segmentation of images. DL-

CONTACT G. Bernabé ✉ gbernabe@um.es 🖹 Computer Engineering Department, University of Murcia, Murcia, Spain

LVTQ has been able to accurately segment the endocardium border and trabeculae and estimate the level of hyper-trabeculation to determine the diagnosis of LVNC. This work has been carried out on 277 patients with hypertrophic cardiomyopathy (HCM).

In this paper, we ask whether this model could be extended to determine the quantification of trabeculae and the diagnosis of LVNC in patients with different types of cardiomyopathies (or multiples cardiomyopathies), which CMR were taken from different scanners and hospitals.

The main contributions of this paper are the following:

- Show in the LVNC diagnosis case the limitations of deep learning models and explain their causes.
- Modify and adapt our previous U-Net CNN architecture to train DL-LVTQ with images collected at different hospitals and several scanners with various cardiomyopathies.
- Validate the new proposal by statistical methods and a new group of medical doctors that endorse our experimental results.

2. Materials and methods

2.1. Hyper-trabeculation quantification in the left ventricle

Due to the continuing discrepancy in the medical community in determining the level of left ventricular trabeculation and the diagnosis of LVNC, we made a semi-automatic proposal called QLVTHC (Bernabé et al. 2017), and an automatic tool called SOST (Bernabé et al. 2018) to help them.

The proposed method aims to segment the different contours of short-axis cardiac MRI in two stages. Firstly, it detects three different shapes of the left ventricle: the compacted external layer, the trabecular zone and the left ventricle cavity. An example with the three contours detected is shown in Figure 1. In the second stage, the area of the trabecular zone and the compacted zone are calculated. Then, the proportion of the trabecular area to the size of the compacted zone for all slices of a patient is obtained. Once all areas of a patient are calculated, the trabecular and compacted volumes are obtained by aggregating the information of all slices.

In this way, the percentage of the trabecular volume to the total volume of the external wall of the left ventricle ($VT\%$) is evaluated according to Equation 1:

$$VT\% = 100 \cdot \frac{Trab.\ volume}{Trab.\ volume + Compacted\ volume} \quad [\%] \quad (1)$$

The computation of the $VT\%$ has been performed by adding up the areas obtained for each of the individual slices, as was already computed previously in Bernabé et al. (2018), taking into account that the slice thickness and interslice spacing are the same for all the slices of a specific patient.

To perform a diagnosis, high values of this metric have been used as a cut-off point or indicator of LVNC (Jacquier et al. 2010; Bernabé et al. 2017, 2018). In fact, in the tool QLVTHC, we proposed a validated threshold of 27.4% to differentiate
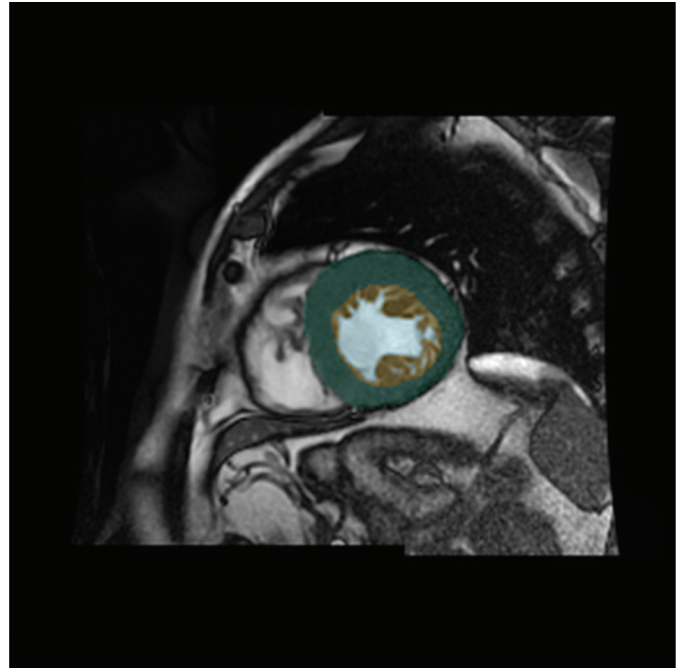
Figure 1. Segmentation of left ventricle, highlighting the external layer (green), the left ventricle cavity (blue) and the trabecular zone or non-compacted area (yellow).

between patients with LVNC and healthy patients (Bernabé et al. 2017).

2.2. The DL-LVTQ approach

In our previous work (Rodríguez-de-Vera et al. 2022), we have selected a 2D U-Net-like CNN based on the well-known U-Net (Ronneberger et al. 2015), from among other networks (Rodríguez-de-Vera et al. 2021), because it provides the best performance to the segmentation of the three contours in the left ventricle.

A U-Net architecture contains a symmetric encoder and decoder pathway, with skip connections between the corresponding layers. The context information is encoded in the down-sampling path and transferred with skip connections to the up-sampling path to decode feature-map and the segmentation maps.

This U-Net architecture was trained on 2D short-axis magnetic resonance imaging to segment the left ventricle's internal cavity, external wall and trabecular tissue, obtaining the following numbers for 25 test images: the average and standard deviation of the Dice coefficient for the internal cavity, external wall and trabeculae were 0.96 ± 0.00, 0.89 ± 0.00 and 0.84 ± 0.00, respectively. Besides, to validate the diagnosis, two cardiologists visually graded the outputs generated automatically by DL-LVTQ according to Gibson's scale (Gibson et al. 2004; Zaid et al. 2008; Bernabé et al. 2015), over a set of 25 patients and 99.5% of the slices were determined without diagnostically relevant issues, that is, the segmentation in internal cavity, external wall and trabecular tissue was visually correct. Note that using Gibson's scale, the outputs generated were visually graded from 1 to 5. A value higher than or equal to 4.0 indicates that the segmentation does not represent diagnostically significant differences.

Table 1. Mean %(± standard deviation across the fivefolds) of the Dice coefficient for the compacted external layer (CEL), the left ventricle cavity (LVC), the trabecular zone (TZ), and the average of the three contours to inference on different populations.

Population	Dice CEL	Dice LVC	Dice TZ	Average Dice
P on P	0.89 ± 0.10	0.94 ± 0.10	0.83 ± 0.15	0.89 ± 0.09
X on P	0.82 ± 0.16	0.92 ± 0.15	0.78 ± 0.19	0.84 ± 0.15
H on P	0.83 ± 0.13	0.88 ± 0.13	0.70 ± 0.21	0.80 ± 0.13

In this paper, we explore the possibility of using the DL-LVTQ approach for diagnosing new images that have not been seen before. Deep learning-based techniques sometimes struggle to extract the essential features of images, which can lead to good accuracy for learning images but lower scores for test images (known as the over-fitting problem). In a previous paper (Rodríguez-de-Vera et al. 2022), we performed a five-fold cross-validation process, dividing the image dataset into five equal-sized folds. One of the folds was used as the test set, while the model was trained on the remaining four folds, which were split into training-validation sets with an 80–20% ratio. It is important to note that the training, validation and test sets were defined based on individual patients. It is essential to note that the training, validation and test sets were defined based on individual patients.

However, we were concerned regarding the dataset, as all the patients (277, which has a mean of 7 images per patient) suffered from the same cardiomyopathy. Therefore, we extended our original working dataset into three populations: the original P group (augmented with 16 new patients, therefore the P group has 293 patients), the X group and the H group (see Section 2.4 for more details on these populations). Table 1 gives our new results on diagnosing (inferencing) different populations P, X and H, showing the mean and standard deviation of the Dice coefficients for the three contours. The model trained individually on the set P inferences patients of populations X and H with a significant drop in the Dice coefficients reaching 0.70 in the trabecular zone for patients belonging to H, causing difficulties in reconstructing some slices. The average Dice value for the different contours decreases to 0.84 and 0.80 for patients of sets X and H, respectively. Therefore, the neural network model training with images from the P populations does not obtain a good generalisation for patients from other populations.

2.3. Neural network architecture

After achieving those results, we looked for reasons that could explain them.

The first point was the image quality. Input images that conform to the ground-truth delineation were stored at hospital computers with low resolution (128 × 128 pixels, 92 × 92 pixels and sometimes could decrease to 64 × 64 pixels) to save storage space on their hard disks. We realised that image size is dragging a fundamental problem that limits the perfect reconstruction of the output images. Trying to overcome this issue, we tuned the QLVTHC tool to generate an output image size of 512 × 512 pixels to avoid an excessive loss of ground-truth

information. Note that we do not apply any other pre-processing to the images.

Afterwards, we elaborated on a better neural network architecture to cope with this problem. Based on our previous research on evaluating different networks (Rodríguez-de-Vera et al. 2021), we decided to continue using a U-Net-like network. Then, we have reconfigured the U-Net in the same way as in (Rodríguez-de-Vera et al. 2022), but the image size of the single-channel inputs is 512 × 512 pixels (to match the QLVTHC tool output). Therefore, the number of steps to encode and decode has been increased adequately.

In each step of the down-sampling path (encode phase), which is formed by seven levels, a convolutional block composed of two 3 × 3 convolutions with batch normalisation and activations with the rectified linear unit (ReLU) is followed by a max-pooling operation with size 2 × 2 and stride of 2. The number of feature maps is doubled in each down-sampling step. A convolution block is used in the bottom layer to link the down-sampling to the up-sampling path, and a bilinear interpolation is implemented in the up-sampling path instead of up-convolution (used in the original U-Net design). Identically to the encoding, in the decode phase (up-sampling path), a series of bilinear interpolation and convolutional blocks are repeated to reach the final convolution layer.

The final convolution layer includes a softmax activation function to obtain the output segmentation maps in four classes: the left ventricle cavity, the external layer, the trabecular zone and the background. The final segmentation maps also have a size of 512 × 512 pixels. Figure 2 shows the architecture diagram used in our research.

For simplicity, we apply a 2D segmentation to the input slices of the network. We apply Z-score standardisation and we perform data augmentation by applying random rotations of 90, 180 and 270 degrees with a 0.25 probability. As a loss function, we use a linear combination of two components: Lovász-Softmax loss ([Berman et al. 2018]) and weighted binary cross-entropy loss. We also apply the Rectified Adam (RAdam) optimiser (Liu et al. 2020) and a fivefold cross-validation based on a threshold of 27.4% previously described in Section 2.1. We made these selections based on our previous work (Rodríguez-de-Vera et al. 2022).

2.4. Populations

We have compounded a dataset consisting of three groups of patients:

(1) The set labelled P has 293 patients diagnosed with hypertrophic cardiomyopathy.
(2) The set labelled X has 58, unclassifiable patients with different cardiomyopathies diagnosed as non-
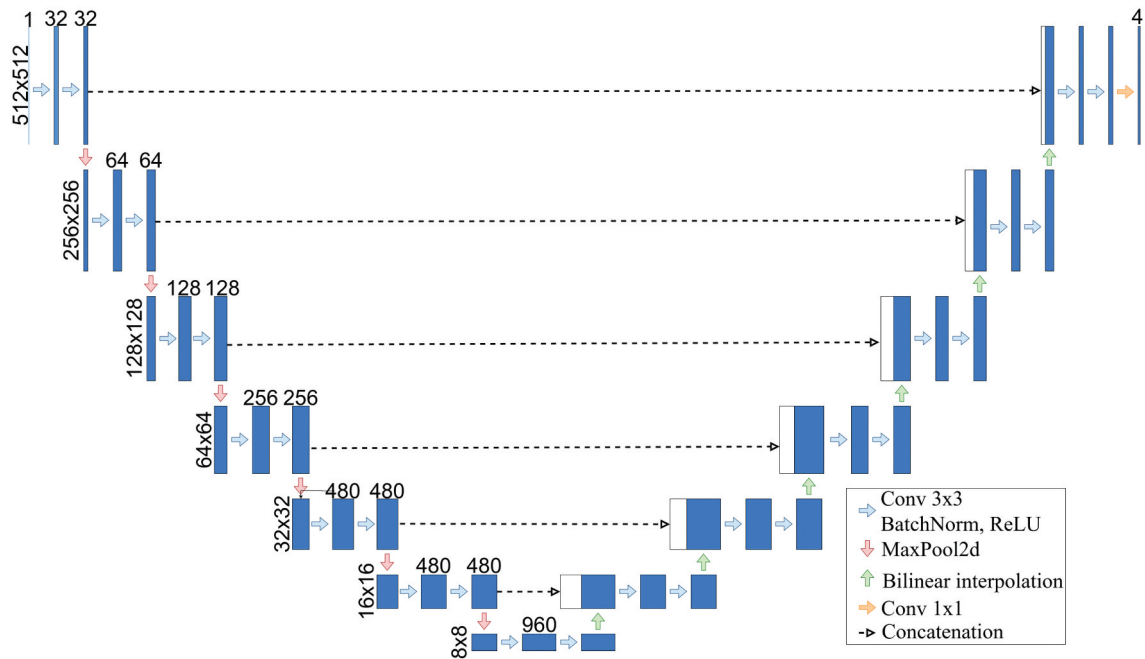
Figure 2. Architecture of the U-Net segmentation network in this research. Each blue box corresponds to a multi-channel feature map.

compacted cardiomyopathy, RV (right ventricular) or LV (left ventricular) arrhythmogenic cardiomyopathy, DCM (dilated cardiomyopathy), HCM (Hypertrophic Cardiomyopathy), i.e. multiple cardiomyopathies, and inherited cardiomyopathies. Therefore, medical specialists classified these patients as having more than one heart disease.

(3) The set labelled H has 28 patients with previously diagnosed LVNC cardiomyopathy meeting Petersen's criteria Petersen et al. (2005).

The population has been selected from an inherited cardiomyopathy clinic. A good-quality MRI in the first test was acquired in the short axis.

The MRI images are obtained at three different hospitals: Virgen de la Arrixaca of Murcia (HVAM), Mesa del Castillo of Murcia (HMCM) and Universitari Vall d'Hebron of Barcelona (HUVHB). HVAM manages two 1.5 T scanners, Philips and General Electric, with acquisition matrices of 256×256 pixels and 224×224 pixels and pixel spacing of $1.5 \times 1.5 \times 8mm$ and $1.75 \times 1.75 \times 8mm$, respectively. HMCM possesses the same General Electric model scanner as HVAM. HUVHB has a 1.5 scanner Avanto of Siemens, where the acquisition matrix is 224×224 pixels. The LV function is determined with balanced steady-state free precision (b-SSFP) sequences, where the

repetition interval is established to 3.8 ms for HMCM and HUVHB, whereas HVAM uses 3.3 ms. Other parameters like echo time, flip angle, echo train length, slice thickness, slice gap and phases are fixed to 1.7 ms, 60°, 23, 8 mm, 2 mm and 20 phases for all scanners. All patients have been monitored in apnoea, in synchronisation with the ECG, and without a contrast agent.

In short, the final dataset comprises 3044 slices of 379 patients. The number of patients with LVNC is 223. The number of slices per patient ranges from 1 to 14; the median and the mean are seven slices per patient.

3. Results and discussion

3.1. Training with all the populations

We started by training our modified neural network described in Section 2.3 with mixed images from different populations. Table 2 describes the obtained inference results when training with images from the P and X populations (P+X in the table) and when training with images from all populations (P+X+H in the table). We report the mean and standard deviation of the Dice coefficients for the three contours and the average Dice, split by patients belonging to populations P, X and H.

Table 2. Mean %($\pm$ standard deviation across the fivefolds) of the Dice coefficient for the compacted external layer (CEL), the left ventricle cavity (LVC), the trabecular zone (TZ) and the average of the three contours to inference on different populations.

Population	Dice CEL	Dice LVC	Dice TZ	Average Dice
P on P+X	0.89 ± 0.10	0.94 ± 0.10	0.84 ± 0.15	0.89 ± 0.09
X on P+X	0.83 ± 0.17	0.93 ± 0.16	0.79 ± 0.21	0.85 ± 0.16
H on P+X	0.84 ± 0.12	0.90 ± 0.12	0.73 ± 0.20	0.82 ± 0.13
P on P+X+H	0.89 ± 0.09	0.94 ± 0.09	0.84 ± 0.14	0.89 ± 0.09
X on P+X+H	0.84 ± 0.14	0.93 ± 0.14	0.80 ± 0.18	0.86 ± 0.13
H on P+X+H	0.86 ± 0.09	0.92 ± 0.10	0.79 ± 0.16	0.86 ± 0.09

Table 3. Mean %($\pm$ standard deviation across the fivefolds) of the Dice coefficient for the compacted external layer (CEL), the left ventricle cavity (LVC), the trabecular zone (TZ) and the average of the three contours for different populations.

Population	Dice CEL	Dice LVC	Dice TZ	Average Dice
P	0.88 $\pm$ 0.00	0.95 $\pm$ 0.00	0.83 $\pm$ 0.00	0.89 $\pm$ 0.01
P + X	0.87 $\pm$ 0.02	0.95 $\pm$ 0.01	0.83 $\pm$ 0.02	0.88 $\pm$ 0.01
P + X + H	0.86 $\pm$ 0.02	0.94 $\pm$ 0.02	0.82 $\pm$ 0.03	0.88 $\pm$ 0.02

As we can see, when the network model is trained with P and X populations (P+X), the Dice coefficients improve concerning those obtained with population P (shown in Table 1). However, inferences for patients in the H group remain a bit lower. For example, the Dice coefficient of the trabecular zone and the average Dice value are fixed at 0.73 and 0.82, respectively. Finally, when the network model is trained with P, X and H populations (P+X+H), the results are much better, reaching an average Dice value of 0.86 to infer both a patient belongs to X or H. Therefore, this whole model, in which several patients with different characteristics of X and H are added to the initial P population, helps to solve and reinforce the model, making it more robust and prepared to infer patients with distinct heart diseases from different hospitals.

Table 3 completes the previous results from our modified neural network. This Table reports the training results for the mean and standard deviation of the Dice coefficients for the compacted external layer, the trabecular zone and the left ventricle cavity when using only images from the P population (P), an image mixed from the P and X population (P+X) and, finally, an image combined from the P, X and H population (P+X+H). As we can see, there is high accuracy in detecting the compacted external layer and the left ventricle cavity, keeping very close values even though new sets of patients, such as X and H, are included. This result shows that it is feasible to increase the dataset so that the network is trained for a broader spectrum of patients with distinct features. It is acquired with various machines and in separate hospitals. In addition, it is essential to note that the standard deviation is very low or close to zero, varying very slightly as new sets of patients are added, showing that the results are pretty robust.

Regarding the Dice coefficient for the trabecular zone, the obtained values are a bit lesser due to the intrinsic difficulty associated with this zone, formed by several separate parts, an aspect also experienced by cardiologists in determining this controversial area. Therefore, we can claim that the network accurately distinguishes the trabecular zone. Moreover, it is essential to remember that there is a limitation in achieving higher values for the Dice coefficients due to the low image size at which MRIs are stored in medical centres.

By using the neural network presented in Section 2.3 and taking as a reference the values obtained by the semi-automatic QLVTHC proposal (Bernabé et al. 2018), we compute the mean error and standard deviation between the $VT\%$ and the values obtained in our modified network model. We report

Table 4. Confusion matrix for the population P by a threshold of 27.4%.

	Reference diagnosis		
	LVNC	No LVNC	Total
LVNC	145	22	167
No LVNC	24	123	147
Total	169	145	314

Table 5. Confusion matrix for the population P+X by a threshold of 27.4%.

	Reference diagnosis		
	LVNC	No LVNC	Total
LVNC	193	10	203
No LVNC	39	117	156
Total	232	127	359

5.01 ± 0.19 mm^3 for the population P, 5.39 ± 0.90 mm^3 for the population P+X and 5.63 ± 1.26 mm^3 for the population P+X +H. Therefore, there is no significant increase in committed errors despite expanding the dataset with patients with different cardiomyopathies.

3.2. Statistical evaluation

From a medical point of view, it is important to evaluate our modified network model statistically. Consequently, the confusion matrices (based on 27.4% threshold validated by QLVTHC) are presented in Tables 4, 5 and 6 for populations P, P+X and P +X+H, respectively. In addition, Table 7 shows the accuracy, sensitivity, specificity and Kappa values.

As we can see, the network's accuracy and Kappa are slightly increased. In contrast, the sensitivity and the specificity are also increased in a more significant way by including the sets X and H. This means we have a better-prepared network for a wide range of cardiomyopathies. This network allows any cardiologist to automatically obtain the inference of any patient without spending considerable and disproportionate time. Moreover, for the complete dataset, the positive and negative

Table 6. Confusion matrix for the population P+X+H by a threshold of 27.4%.

	Reference diagnosis		
	LVNC	No LVNC	Total
LVNC	210	13	223
No LVNC	34	122	156
Total	244	135	379

Table 7. Accuracy, sensitivity, specificity and Kappa values for different populations.

Population	Accuracy	Sensitivity	Specificity	Kappa
P	0.85	0.87	.85	.71
P + X	0.86	0.95	.92	.72
P + X + H	0.88	0.94	.90	.74

Table 8. Confusion matrix for the population P+X+H by a threshold of 27.1%.

	Reference diagnosis		
	LVNC	No LVNC	Total
LVNC	215	8	223
No LVNC	36	120	156
Total	215	128	379

predictive values achieve 0.94 (210/223) and 0.78 (122/156), respectively, improving the detection of LVNC patients and maintaining healthy patients against our previous proposal (Rodríguez-de-Vera et al. 2022).

Since the cut-off point of 27.4% was validated using the QLVTHC tool, it is possible that a better classification can be achieved by varying the threshold point using DL-QLVT. Hence, it is necessary to calculate receiver operating characteristics (ROC) curve analysis of the $VT\%$ obtained from the modified network to determine that the optimal cut-off point is 27.1%. The corresponding confusion matrix can be found in Table 8. The area under the ROC curve is 0.94 (95% confidence interval, 0.91–0.96), and with a threshold of 27.1%, the accuracy, sensitivity, specificity and Kappa values are 0.88, 0.96, 0.94 and 0.75, respectively. Therefore, the positive and negative predictive values achieve 0.96 (215/223) and 0.77 (120/156), respectively, further improving the detection of LVNC patients and very slightly worsening the detection of healthy patients.

3.3. Medical validation

We have conducted an evaluation of the outputs that were automatically generated by the updated DL-LVTQ. This evaluation was carried out by two cardiologists who work in different hospitals. Our aim was not to introduce any subjective assessment of the images, as it would contradict the purpose of the proposed automatic tool. However, it is essential that different cardiologists who were not involved in collecting the images can detect any errors.

To be more specific, these two cardiologists have scored all slices of patients belonging to population H. They have determined that 98.9% of the slices of H-patients are entirely valid from a medical point of view, with a score equal to or higher than 3.5, or without significant differences to make a diagnosis.

For example, Figure 3 shows the segmentations (green indicates the compacted external layer of the left ventricle cavity and yellow the trabecular zone) for one patient, where two cardiologists scored a 5 (exact match) or 4.5 on the
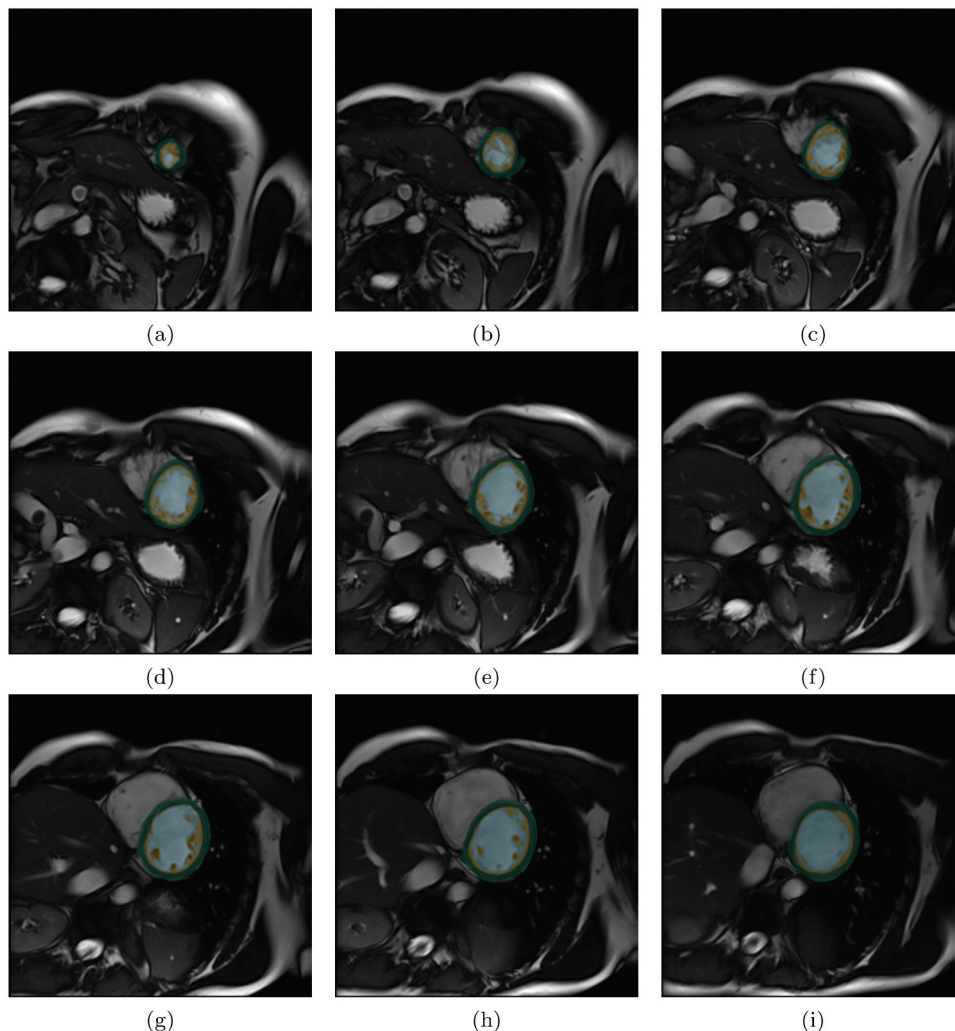
Figure 3. Output slices for the patient H15. Green indicates the compacted external layer of the left ventricle cavity, and yellow is the trabecular zone.

different slices. However, fewer slices of H patients have not been rebuilt correctly, so it could be advisable to increase the number of such patients in the dataset, causing a double benefit that will imply an improvement of the Dice coefficient for the different contours and better image reconstruction.

The scores given by two different cardiologists showed a mean difference of 0.58 ± 0.51 ($13.1\% \pm 11.8\%$) in terms of inter-observer variability. This discrepancy is due to the subjectivity of the cardiologists from different hospitals. The main source of disagreement is observed in slices near the apical or basal ends, where the accuracy of the network is lower due to the small size of apical-end slices or the difficulty in delimiting borders in the adjacent cavities for basal-end slices. This fact has been reported previously in Rodríguez-de-Vera et al. (2022) and Bernard et al. (2018).

4. Conclusions and future work

In a previous paper, we presented DL-LVTQ (Rodríguez-de-Vera et al. 2022). DL-LVTQ was a deep-learning left ventricular trabecular quantification approach based on a U-Net CNN architecture. DL-LVTQ was ready to process and determine the presence of LVNC in patients with hypertrophic cardiomyopathy.

From a medical point of view, the results generated by DL-LVTQ do not present notable differences for making a diagnosis in most cases, which favours inter-hospital collaboration and research in difficult-to treat patients with multiple cardiomyopathies.

In this research paper, we have made modifications to the U-Net and customised it to handle the unique characteristics of three groups of patients who suffer from different types of cardiomyopathies simultaneously and inherited cardiomyopathies. The images used in this study were obtained from various hospitals, captured by multiple scanners and adjusted to fit into the new data set.

Our modified network model shows remarkable accuracy in segmenting the endocardium border and trabeculae, and it performs better than the previous manual, semi-automatic and automatic proposals with robustness and speed. This means that cardiologists can obtain the inference of any patient automatically, without significant effort and personal bias. Additionally, we can assert that the system is equipped and adaptable to handle a diverse group of patients.

In summary, using deep learning techniques to automatically diagnose LNVC cardiomyopathy from MRI images is a complex and challenging process. However, we believe that our results are promising and worthy. While there is still room for improvement, an automated diagnosis system will provide cardiologists with a fast and reliable way to determine LVNC without spending a significant amount of time. This will help eliminate human error and subjectivity, and facilitate the evolution of LVNC and other heart diseases. Collaborations between hospitals and research groups will enable the diagnosis of unclassified patients with multiple cardiomyopathies. Ultimately, this approach will benefit the patients the most.

Notes on contributors

Gregorio Bernabé received the M.S. and Ph.D. degrees in Computer Science from the University of Murcia (Spain) in 1997 and 2004, respectively. In 1998, he joined the Computer Engineering Department of the University of Murcia and became an Associate Professor in May 2004. He has developed several courses on Computer Structure and Computer Architecture. He is currently working on two main research lines: Heterogeneous computing using CMP architectures, GPUs, and accelerators (XPUs), and deep neural network learning process using HPC techniques to improve the performance of medical applications. He has published more than 50 refereed papers in different journals and conferences, including JCR journals such as Journal of Supercomputing, Journal of Computer Methods and Programs in Biomedicine, International Journal of Parallel Programming and Journal of Computational Science.

Pilar González-Férez is Associate Professor in the area of Computer Architecture and Technology at the Faculty of Computer Science of the University of Murcia. She obtained her PhD in Computer Science in 2012 at the same University of Murcia. Her research focuses on diverse areas such as operating systems, distributed systems, key-value database design, education research and deep neural networks. She has participated in several research projects and published her research results both in journals and conferences.

José M. García is a Professor of Computer Architecture at the University of Murcia (Spain) and the founding member and head of the Research Group on Parallel Architecture and Computing (GACOP). His research interest has always focused on the topic of Supercomputing. It started within the line of the interconnection network for high-performance computing systems. Subsequently, research work was carried out on issues of improving the performance of single-chip multiprocessors (CMPs), paying particular attention to the internal architecture of the chip, the coherence of the caches, and virtualization. He is currently working on two main research lines: Heterogeneous computing using CMP architectures, GPUs, and accelerators (XPUs), and Acceleration of bio-inspired algorithms, bioinformatics algorithms, and deep neural network learning process using HPC techniques. He has directed eighteen Doctoral Theses and has published more than 150 refereed articles in international journals and conferences. Prof. García is a member of HiPEAC, the European Network of Excellence in Architecture and High-Performance Integrated Compilation. Furthermore, he is also a member of various international associations such as the IEEE and ACM.

Guillem Casas studied Medicine Degree at the Universitat de Barcelona (UB) from 2006 to 2012 and he then completed his 5-year Cardiology Residency training at Hospital Universitari Vall dHebron (HUVH, Barcelona, Spain) from 2013 to 2018. Since 2018, he is a consultant cardiologist of the cardiovascular imaging department and of the inherited cardiac diseases unit at HUVH, where he has been carrying both clinical and research duties. He is currently undertaking his PhD studies at Universitat Autònoma de Barcelona (UAB) on left ventricular noncompaction – excessive trabeculation of the left ventricle. He has published a number of high-impact articles, participates in several multicentric studies and has received several competitive research grants.

Josefa González-Carrillo is a Cardiology Specialist Physician at Virgen de la Arrixaca University Hospital since 2003. Her training as Cardiologist (2003) is completed with a sub specialization in cardiac imaging accredited by the Spanish Society of Cardiology (2013) and by the European Society of Cardiology (2012). She also has level 2 accreditation from the German Society of Cardiology to perform and interpret Cardiac Resonance studies (2007). Doctorate from the University of Murcia in 2015 with the thesis: "Non-compaction cardiomyopathy, quantification and 2 prognoses". She has collaborated with the Computer Engineering Department of the University of Murcia in the development of software for image diagnosis in cardiology. She has also participated in several multicentre research projects through the performance and interpretation of echocardiography and cardiac MRI.

ORCID

G. Bernabé http://orcid.org/0000-0002-7265-3508
P. González-Férez http://orcid.org/0000-0003-1681-5442
J. M. García http://orcid.org/0000-0002-6388-2835
G. Casas http://orcid.org/0000-0001-7122-320X
J. González-Carrillo http://orcid.org/0000-0003-1981-1885

Data availability statement

References

Arbustini E, Weidemann F, Hall JL. 2014. Left ventricular noncompaction: a distinct cardiomyopathy or a trait shared by different cardiac diseases? J Am Coll Cardiol. 64(17):1840–1850. doi: 10.1016/j.jacc.2014.08.030.

Bartoli A, Fournel J, Bentatou Z, GH G, Lalande A, Bernard M, Boussel L, Pontana F, Dacher J, Ghattas B, et al. 2020. Deep learning-based automated segmentation of left ventricular trabeculations and myocardium on cardiac MR images: a feasibility study. Radiolo Art Int. 25(3):–. doi:10.1148/ryai.2020200021.

Berman M, Triki AR, Blaschko MB 2018. The Lovasz-softmax loss: a tractable surrogate for the optimization of the intersection-over-union measure in neural networks. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition; Salt Lake City, UT, USA. p. 4413–4421. doi:10.1109/CVPR.2018.00464.

Bernabé G, Casanova JD, Casas G, González-Carrillo J 2020. A highly accurate method for quantifying LVNC cardiomyophaty. AMIA 2020 Annual Symposium; Chicago, USA. p. 223–232.

Bernabé G, Casanova JD, Cuenca J, González-Carrillo J. 2018. A self-optimized software tool for quantifying the degree of left ventricle hyper-trabeculation. J Supercomput. 75(3):1625–1640. doi: 10.1007/s11227-018-2722-x.

Bernabé G, Casanova JD, González-Carrillo J, Gimeno-Blanes JR. 2021. Towards an enhanced tool for quantifying the degree of lv hyper-trabeculation. J Clin Med. 10(3):–. doi: 10.3390/jcm10030503.

Bernabé G, Cuenca J, Giménez D, de Teruel PE L, González-Carrillo J. 2015. A software tool for the automatic quantification of the left ventricle myocardium hyper-trabeculation degree. Procedia Comput Sci. 51:610–619. doi: 10.1016/j.procs.2015.05.329.

Bernabé G, González-Carrillo J, Cuenca J, Rodríguez D, Saura D, Gimeno JR. 2017. Performance of a new software tool for automatic quantification of left ventricular trabeculations. Revista Española de Cardiología (English Edition). 70(5):405–407. doi: 10.1016/j.rec.2016.07.006.

Bernard O, Lalande A, Zotti C, Cervenansky F, Yang X, Heng PA, Cetin I, Lekadir K, Camara O, Gonzalez Ballester MA, et al. 2018. Deep learning techniques for automatic mri cardiac multi-structures segmentation and diagnosis: is the problem solved? IEEE Trans Med Imaging. 37 (11):2514–2525. doi:10.1109/TMI.2018.2837502.

Biagini E, Ragni L, Ferlito M, Pasquale F, Lofiego C, Leone O, Rocchi G, Perugini E, Zagnoni S, Branzi A, et al. 2006. Different types of cardiomyopathy associated with isolated ventricular noncompaction. Am J Cardiol. 98(6):821–824. doi:10.1016/j.amjcard.2006.04.021.

Captur G, Lopes LR, Patel V, Li C, Bassett P, Syrris P, Sado DM, Maestrini V, Mohun TJ, McKenna WJ, et al. 2014. Abnormal cardiac formation in hypertrophic cardiomyopathy. Circ Cardiovasc Genet. 7(3):241–248. doi:10.1161/CIRCGENETICS.113.000362.

Captur G, Muthurangu V, Cook C, Flett AS, Wilson R, Barison A, Sado DM, Anderson S, McKenna WJ, Mohun TJ, et al. 2013. Quantification of left ventricular trabeculae using fractal analysis. J Cardiovasc Magn Reson. 15(1):36. doi:10.1186/1532-429X-15-36.

Chen C, Qin C, Qiu H, Tarroni G, Duan J, Bai W, Rueckert D. 2020. Deep learning for cardiac image segmentation: a review. Front Cardiovasc Med. 7:–. doi: 10.3389/fcvm.2020.00025.

Choi Y, Kim SM, Lee SC, Chang SA, Jang SY, Choe YH. 2016. Quantification of left ventricular trabeculae using cardiovascular magnetic resonance for the diagnosis of left ventricular non-compaction: evaluation of trabecular volume and refined semi-quantitative criteria. J Cardiovasc Magn Reson. 18(1):24–24. doi: 10.1186/s12968-016-0245-2.

Gibson D, Spann M, Woolley S. 2004. A wavelet-based region of interest encoder for the compression of angiogram video sequences. IEEE Trans Inf Technol Biomed. 8(2):103–113. doi: 10.1109/TITB.2004.826722.

Jacquier A, Thuny F, Jop B, Giorgi R, Cohen F, Gaubert JY, Vidal V, Bartoli JM, Habib G, Moulin G. 2010. Measurement of trabeculated left ventricular mass using cardiac magnetic resonance imaging in the diagnosis of left ventricular non-compaction. Eur Heart J. 31(9):1098–1104. doi: 10.1093/eurheartj/ehp595.

Li C, Song X, Zhao H, Feng L, Hu T, Zhang Y, Jiang J, Wang J, Xiang J, Sun Y. 2021. An 8-layer residual U-Net with deep supervision for segmentation of the left ventricle in cardiac ct angiography. Comput Methods Programs Biomed. 200:105876. doi: 10.1016/j.cmpb.2020.105876.

Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, van der Laak JA, van Ginneken B, Sánchez CI. 2017. A survey on deep learning in medical image analysis. Med Image Anal. 42:60–88. doi: 10.1016/j.media.2017.07.005.

Liu L, Jiang H, He P, Chen W, Liu X, Gao J, Han J 2020. On the variance of the adaptive learning rate and beyond. In: 8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020. Available from: https://openreview.net/forum?id=rkgz2aEKDr.

Mc Namara K, Alzubaidi H, Jackson J. 2019. cardiovascular disease as a leading cause of death: how are pharmacists getting involved?. IPRP. 8:1–11. doi: 10.2147/IPRP.S133088.

Penso M, Moccia S, Scafuri S, Muscogiuri G, Pontone G, Pepi M, Caiani E. 2021. Automated left and right ventricular chamber segmentation in cardiac magnetic resonance images using dense fully convolutional neural network. Comput Methods Programs Biomed. 204:106059. doi: 10.1016/j.cmpb.2021.106059.

Pérez-Pelegrí M, Monmeneu JV, López-Lereu MP, Pérez-Pelegrí L, Maceira AM, Bodí V, Moratal D. 2021. Automatic left ventricle volume calculation with explainability through a deep learning weak-supervision methodology. Comput Methods Programs Biomed Available from. 208:106275. https://www.sciencedirect.com/science/article/pii/S0169260721003497

Petersen SE, Selvanayagam JB, Wiesmann F, Robson MD, Francis JM, Anderson RH, Watkins H, Neubauer S. 2005. Left Ventricular Non-Compaction: Insights From Cardiovascular Magnetic Resonance Imaging. J Am Coll Cardiol. 46(1):101–105. doi: 10.1016/j.jacc.2005.03.045.

Rodríguez-de-Vera JM, Bernabé G, García JM, Daniel S, González-Carrillo J. 2022. Left ventricular non-compaction cardiomyopathy automatic diagnosis using a deep learning approach. Comput Methods Programs Biomed Available from. 214:106548. https://www.sciencedirect.com/science/article/pii/S0169260721006222

Rodríguez-de-Vera JM, González-Carrillo J, García JM, Bernabé G. 2021. Deploying deep learning approaches to left ventricular non-compaction measurement. J Supercomput. 77(9):10138–10151. doi: 10.1007/s11227-021-03664-0.

Ronneberger O, Fischer P, Brox T. 2015. U-net: convolutional networks for biomedical image segmentation. Lecture Notes In Computer Science

Medical Image Computing And Computer-Assisted Intervention – MICCAI. 9351:234–241. doi:10.1007/978-3-319-24574-4_28.

Sze V, Chen YH, Yang TJ, Emer JS. 2017. Efficient processing of deep neural networks: a tutorial and survey. Proc IEEE. 105(12):2295–2329. doi: 10.1109/JPROC.2017.2761740.

Towbin JA, Lorts A, Jefferies JL. 2015. Left ventricular non-compaction cardiomyopathy. Lancet. 386(9995):813–825. doi: 10.1016/S0140-6736(14)61282-4.

Udeoji DU, Philip KJ, Morrissey RP, Phan A, Schwarz ER. 2013. Left ventricular noncompaction cardiomyopathy: updated review. Ther Adv Cardiovasc Dis. 7(5):260–273. PMID: 24132556. doi: 10.1177/1753944713504639.

World Health Organization. 2022. Cardiovascular diseases. Accessed Feb 2 2024 Fri 14:43:26. https://www.Who.int/health-Topics/cardiovascular-Diseases.

Zaid AO, Abdessalem S, Mourali MS, Farhati A, Bouallègue A, Mechmeche R, Olivier C 2008. Coronary angiogram video compression adapted to medical imaging applications. Annual International Conference of the IEEE Engineering in Medicine and Biology Society Vancouver, Canada. p. 410–413. doi:10.1109/IEMBS.2008.4649177.