A U-Net with Statistical Shape Restrictions Applied to the Segmentation of the Left Ventricle in Echocardiographic Images

U-Net con Restricciones Estadísticas de Forma, Aplicada a la Segmentación del Ventrículo Izquierdo en Imágenes de Ecocardiograma

Eduardo Galicia-Gómez[1] , Fabián Torres-Robles[2] , Jorge Pérez-González[3] , Boris Escalante-Ramirez[4]
Fernando Arámbula Cosío[3]

[1]Universidad Nacional Autónoma de México, Posgrado en Ciencia e Ingeniería de la Computación, Ciudad de México - México

[2]Universidad Nacional Autónoma de México, Laboratorio de Física Medica, Instituto de Física, Ciudad de México - México

[3] Universidad Nacional Autónoma de México, Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, Yucatán - México

[4] Universidad Nacional Autónoma de México, Facultad de ingeniería, Ciudad de México - México

ABSTRACT

This paper aims to introduce an innovative approach to semantic segmentation by leveraging a convolutional neural network (CNN) for predicting the shape and pose parameters of the left ventricle (LV). Our approach involves a modified U-Net architecture with a regression layer as the final stage, as opposed to the traditional classification layer. This modification allows us to predict all the shape and pose parameters of a statistical shape model, including rotation, translation, scale, and deformation. The adapted U-Net is trained using data from a point distribution model (PDM) of the LV. The experimental results demonstrate a mean Dice coefficient of 0.82 on good quality images, and 0.66 including mean and low-quality images. Our approach successfully overcomes a common issue encountered in CNN-based semantic segmentation. Unlike the inaccurate pixel classification that often leads to unwanted blobs, our CNN generates statistically valid shapes. These shapes hold significant potential in initializing other methods, such as active shape models (ASMs). Our novel CNN-based approach provides a novel solution for semantic segmentation, offering shapes and pose parameters that can enhance the accuracy and reliability of subsequent medical image analysis methods.

KEYWORDS: convolutional neural networks, echocardiography, left ventricle segmentation, statistical shape analysis

RESUMEN

Este artículo tiene como objetivo introducir un enfoque innovador para la segmentación semántica utilizando una red neuronal convolucional (CNN) para predecir los parámetros de forma y posición del ventrículo izquierdo (VI). Nuestro enfoque implica una arquitectura U-Net modificada con una capa de regresión como etapa final, en contraposición a la capa de clasificación tradicional. Esta modificación nos permite predecir todos los parámetros de un modelo estadístico de formas que incluyen rotación, traslación, escala y deformación. La red convolucional se entrena utilizando datos de un modelo de distribución de puntos (PDM) del VI. Los resultados experimentales muestran un coeficiente Dice promedio de 0.82 para imágenes de buena calidad y de 0.66 cuando se incluyen imágenes de calidad media y baja. Nuestro enfoque supera con éxito un problema común en la segmentación semántica basada en CNNs. A diferencia de la clasificación inexacta de píxeles que a menudo conduce a elementos no deseados (blobs), nuestra CNN genera formas estadísticamente válidas. Estas formas tienen un gran potencial para inicializar otros métodos, como los modelos de forma activa (ASMs). En resumen, nuestro enfoque basado en CNN proporciona una solución innovadora para la segmentación semántica, ofreciendo formas y parámetros de posición que pueden mejorar la precisión y confiabilidad de otros métodos de análisis del VI.

PALABRAS CLAVE: análisis estadístico de forma, ecocardiografía, redes neuronales convolucionales, segmentación del ventrículo izquierdo

Corresponding author

TO: Fernando Arámbula Cosío

INSTITUTION: Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, Unidad Yucatán, Universidad Nacional Autónoma de México

ADRESS: Carretara Mérida-Tetiz Km. 4.5, Ucú, Yucatán, México C.P. 97357

EMAIL: fernando.arambula@iimas.unam.mx

INTRODUCTION

Left ventricle (LV) segmentation in ultrasound images is a crucial process for evaluating the ejection fraction and assessing heart condition. Ultrasound offers real-time imaging with moderate costs and no ionizing radiation. However, image quality is affected by speckle noise and acoustic shadows of adjacent structures. For these reasons, several automatic and non-automatic methods have been developed to perform the segmentation of the LV in ultrasound images.

Some of the main approaches to LV segmentation have been active contours, such as geodesic models [1] and level sets [2], deformable templates, such as shape models [3] registration-based methods [4], and supervised learning models such as data base guided segmentation [5], and hybrid active appearance models [6]. International challenges have contributed to the publication of large image databases that enable the training and testing of supervised learning models [7]. Machine learning models have shown high performance in the segmentation of the LV in echocardiography in 2D and 3D, taking advantage of the public image databases. An example of a machine learning method for LV segmentation is reported in [8]. A structured random forest was developed for automatic segmentation of the myocardium and the LV on an echocardiographic data set of 250 patients, including healthy athletes and cardiac patients. The random forest showed improved segmentation results when compared against an active appearance model.

Statistical shape models (SSMs) and active shape models (ASMs) [9] have been extensively used, with continuous improvements. An enhanced ASM, as reported in [10], incorporates an adaptive strategy to construct appearance models of each landmark point, optimizing the number of principal components (PCs) using mean squared eigenvalue error (MSEE). Preprocessing with a Nakagami filter further improved segmentation results on good quality images from the CAMUS dataset. However, details on the ASM's initial-ization method were not provided in the study.

Recently deep learning methods have been applied with good results for the segmentation of the LV. In [11] is presented a review of the application of convolutional neural networks (CNNs) to the segmentation of the LV in ultrasound and MRI. Fully convolutional [12] and the U-Net architecture [13][14][15] have been successfully applied to the classification of the pixels corresponding to the LV; Different SSMs such as morphological models, snake models, and ASMs have been combined with CNNs to create hybrid architectures in different studies. In [16] is reported a hybrid method with a fast region-based CNN and an ASM for LV segmentation in ultrasound images where adaptive anisotropic diffusion filtering is applied to all images, the fast CNN detects the bounding box around the LV for ASM initialization and the ASM finds the final boundary of the LV. On the other hand, new CNN architectures have been developed for the adjustment of a shape contour. In [17] is reported a new approach for left ventricle segmentation based on a CNN that first, detects 3 landmarks: apex, starting and, end points of the endocardium; Afterwards, a triangle (start-apex-end point) is used as initialization for a deep-snake [18] which is adjusted to the endocardium using circular convolution, good results were obtained in the HMC-QU echocardiography data set. Also, convolutional neural networks based on autoencoders such as U-Net [19], have shown good results when performing segmentation with reduced training sets; however, this type of network has a great disadvantage as they are based on semantic segmentation, they produce areas (blobs) with an inaccurate classification that are sometimes far from the organ to be segmented (see Figure 1).

In this work, is reported a modified U-Net based architecture, which incorporates expert shape knowledge of the LV in a point distribution model (PDM) [9], this approach diverges from the conventional pixel by pixel classification method discussed in [19]. Our model focuses on generating statistically valid shapes for the

left ventricle (LV). This alternative segmentation strategy avoids the production of blobs due to pixel classification errors. Our work contributes to improve the reliability of LV segmentation. The training error of the proposed CNN was calculated as the RSME between the expert pose and shape values and the CNN output at each epoch. During the test stage our U-Net model optimizes the shape and pose parameters corresponding to a non-training ultrasound image of the LV. In the following sections are presented all the details of the PDM of the LV, the U-Net architecture developed and its training parameters. In section III are reported the tests and results on the CAMUS [20] and EchoNet-Dynamic [21] databases. In section IV are presented the discussion and conclusions.
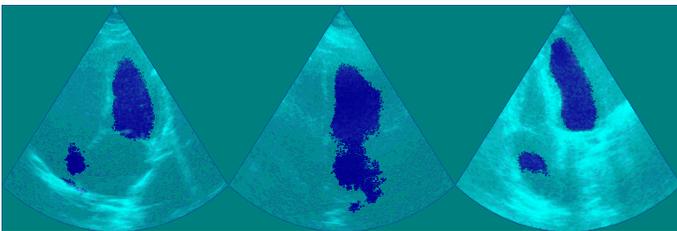

FIGURE 1. **Example of inaccurate classification by the U-Net generating blobs that are not part of the left ventricle.**

MATERIALS AND METHODS

This section presents the proposed methodology. As can be observed in Figure 2, the methodology starts with the training of a U-Net convolutional network, which is fed with a set of echocardiography images (Figure 2a-2b). The target is to estimate the parameters of translation, rotation, scale, as well as a deformation vector (Figure 2c), which are useful to adjust the mean shape obtained from the PDM, and thus perform the segmentation of the left ventricle in the systole and diastole phases (Figure 2d). Finally, the proposed methodology is validated using the Dice coefficient and the Haussdorff distance. The details of each stage are explained in the following subsections.

Point distribution models

Point distribution models (PDMs) have been widely

used for modeling complex structures, such as the organs of the human body [9][22] and specifically in this paper the left ventricle. The model is based on a set of points known as *"landmarks"*, which must be a fixed number and should correspond to the same position along the contour of each example in the training set, as shown in Figure 3.
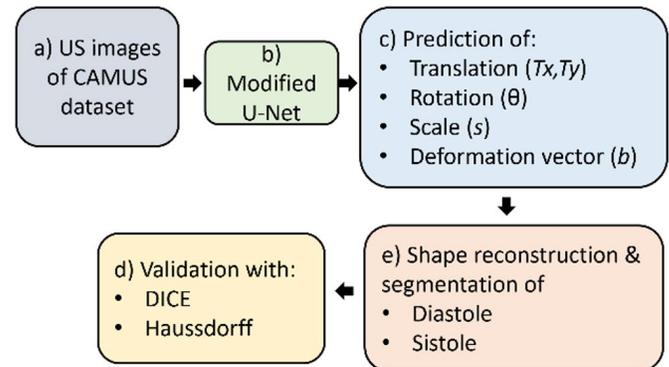

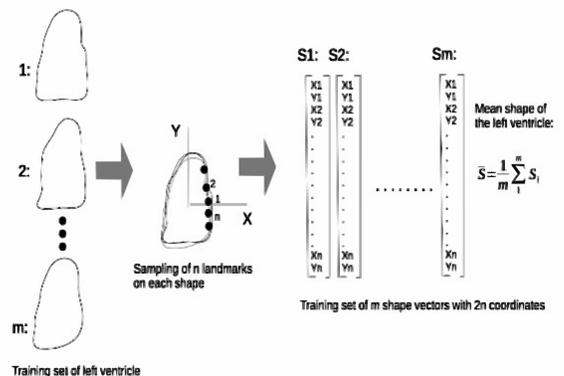**FIGURE 2.** **Workflow of the proposed method.**

FIGURE 3. **Landmark sampling of each left ventricle shape in the training set.**

A shape vector is constructed by concatenation of all the $[x,y]$ coordinates of the landmarks of each training shape. Subsequently, with this set of landmarks, the mean shape ($\bar{x}$) is calculated as proposed by Cootes *et al.*, [9], and the training set is aligned following the Procrustes algorithm [23]. Finally, a principal component analysis (PCA) is performed on the covariance matrix of this set, obtaining the shape parameters (P,b) according to the 90 % of the explained variance, thus enabling the generation of new instances of the left ventricle, as described in Equation (1).

$$\hat{x} = \overline{x} + Pb \qquad (1)$$

Where:

$\hat{x} = new\ shape.$

$\overline{x} = mean\ shape.$

$P = principal\ eigenvector\ matrix.$

$b = deformation\ parameter\ vector.$

Shape and pose parameters.

In this work we constructed a PDM of the left ventricle with the purpose of modeling the deformation of each shape in the training set. The deformation of the PDM for each left ventricle shape can be calculated by solving for b in Equation 1, which yields Equation 2, where b contains the *shape parameters* of each training shape. The *pose parameters* (rotation, translation in X, translation in Y and scale) were obtained as follows: for each of the contours in the training set the rotation and scale were computed as suggested in [9] while translation in X and Y were calculated as the average value in each axis, on each training example.

$$b = (\hat{x} - \overline{x}) * P' \qquad (2)$$

U-Net with statistical shape restrictions

We adapted a convolutional neural network architecture that uses the features extracted by the convolutional filters to predict the values of *shape and pose* of the PDM of left ventricle in this way, all segmentations are *valid shapes* restricted by the training of the network, thus avoiding the blobs shown in Figure 1. The ultimate goal of this CNN is to predict the values of *rotation, translation,* scale and the deformation vector b corresponding to a given ultrasound image of the left ventricle.

Network architecture

To provide a better understanding of the proposed U-Net architecture (see Figure 4), it will be divided into 4 blocks:

1. Input block: Is an image input layer of size 256x256x1 of the form: (width, height, channels).

2. Encoder - Decoder block: The encoder block is formed by convolutional filters of size 3x3 and its goal is to extract relevant features of the image, also the encoder path reduces the spatial resolution of the extracted feature maps applying maxpooling operations. On the other hand, the decoder path upsamples the features maps and preserves the spatial resolution of the input while also performs convolutional operations. By employing the skip connections from the encoder, the decoder layers enhance their ability to detect and fine-tune the features within the image as explained in [19].

3. Fully connected block: It consists of a flatten layer followed by a set of fully connected layers. The purpose of this block is to link the features extracted by the *encoder-decoder* block with the shape and pose parameters related to the input image.

4. Output block: Finally at the end of the fully connected block is a regression layer that allows to predict the desired parameter: rotation, translation, scale, or *b*-vector.

Network Training

Using the pose and shape parameters described above, an input vector was constructed for each training image I_i as follows in Equation (3):

$$V_i = [I_i\ \theta_i, x_i, y_i, s_i, b_i] \quad (3)$$

Where:

$I_i = the\ i - th\ image\ of\ the\ training\ set.$

θ_i = the i-th rotation value.

x_i = the i-th x translation value.

y_i = the i-th y translation value.

s_i = the i-th scale value.

b_i = the i-th deformation vector.

The vector V_i then becomes the input of the proposed CNN, the intention of training a U-Net architecture is to take full advantage of its ability to extract features throughout the encoder stage, the idea is then to use these features to predict the values of pose and shape of the left ventricle, assuming that each LV image corresponds to a single contour and these are never the same between patients. Also the network *loss* during the training stage is calculated as the RMSE between the values of shape and pose predicted by the network (as output from the regression layer) and the *ground truth* contained in each training vector V_i through an iterative process called: stochastic gradient descent (SGD) the network adjusts its weights according to the values of the loss trying to decrease it in each iteration, the lower the loss the better the prediction of the pose and shape values.

Predicted shape reconstruction

Once the network has been trained, it is possible to predict the values of θ,Tx,Ty,*s* and *b* following the diagram in Figure 2. These values allow reconstructing the LV contour corresponding to the input image, first by applying *b* and *P* to the mean shape $(\bar{x})$ to obtain the deformation of the left ventricle (see Equation 2) and then by multiplying the resulting shape $(\hat{x})$ by the transformation matrix in Equation 4, thus obtaining the reconstructed LV shape *(S)*. Finally, *S* is placed in the input image to obtain the corresponding LV segmentation.

$$S = \hat{x}s \begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix} + [Tx, Ty] \quad (4)$$

Where:

$\theta = the\ rotation\ predicted\ by\ theCNN.$

$s = the\ scale\ predicted\ by\ theCNN.$

$[Tx, Ty] = two - column\ translation$

$values.$

Validation

To validate the proposed method, we used two measures:

1. Dice Coefficient: This method allows us to determine the area of overlap between the segmentation mask of the expert and the segmentation mask of our method. A higher overlap area corresponds to a higher Dice coefficient value. Therefore, the larger the coefficient value, the better the segmentation. This coefficient is within the range of 0 to 1.

2. Haussdorff Distance: It measures the distance between two sets of points. In this case, it measures the distance between the contour marked by the expert and the contour reconstructed with the network values. A smaller Haussdorff distance indicates a greater similarity between the two contours.

RESULTS AND DISCUSSION

This section presents how the training and test dataset were created, the CNN training parameters and the segmentation results obtained by the CNN reported in this paper. It is also important to mention that the data extraction and image preprocessing were performed in Python 3.7, while the development, training and testing of the proposed network were done using MATLAB 2022b.

Training dataset

A set of 800 images (400 systole and 400 diastole) were randomly selected from the CAMUS database [20], this set includes images of good, regular, and poor quality, then, data augmentation modifying rotation, translation, scale, brightness, and contrast was applied to achieve a total of 4800 training images an example of these images is shown in Figure 5. With the use of data augmentation, we managed to attain greater diversity in the training set and, consequently, enhance the generalization achieved by the network.
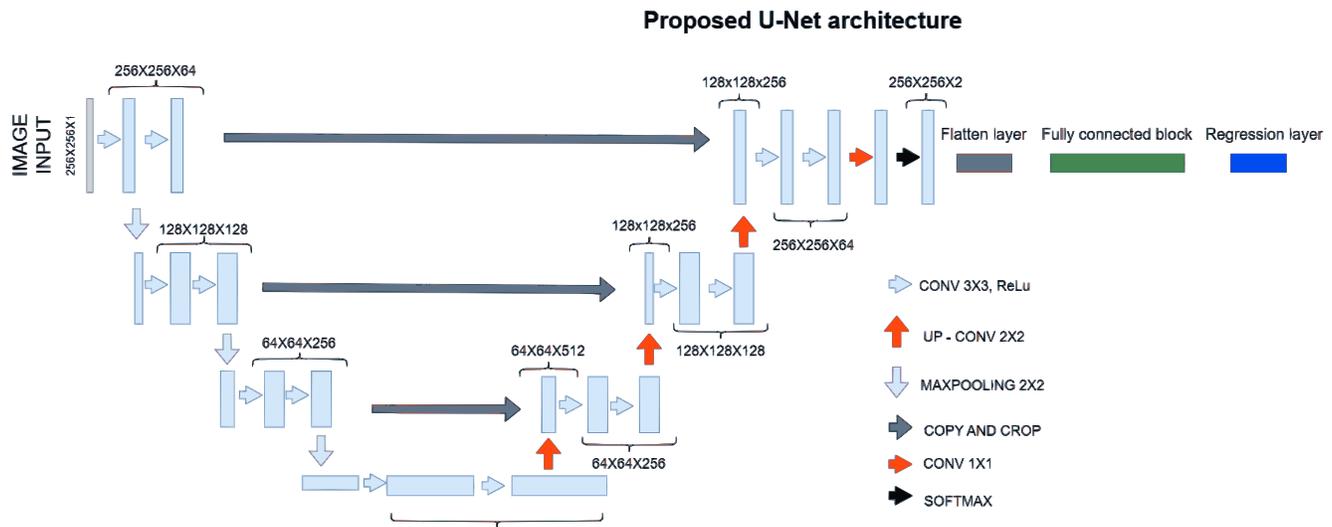
FIGURE 4. Proposed U-Net architecture.

Test dataset

The test set used to evaluate our method was divided into 2 parts: the first consists of 98 images from the CAMUS database, comprising 49 systole images and 49 diastole images (hold-out set). The second part consists of non-training images from the EchoNet-dynamic database [21], these images were extracted from .AVI videos at the end of the systolic and diastolic cycles. The task was done by Cervantes – Guzmán in [24]. The above-mentioned images were subsequently anno-tated by an expert and the segmentation masks were obtained. The total number of systole images is 207, while for diastole, it is 211. It is worth noting that the resolution of the images from the EchoNet-dynamic database is 112x112 pixels, so they were scaled to 256x256, the resolution at which the proposed network is trained.
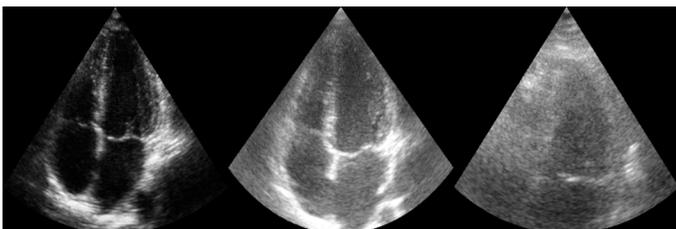

FIGURE 5. Examples of good, regular and bad quality images from the CAMUS database.

CNN training parameters

The training parameters for the CNN were defined as follows:

- Training Images = 4560 images.
- Validation images = 240 images.
- Number of epochs = 20.
- Learning rate = 0.0001.
- Batch size = 32.
- Encoder depth = 3.
- Start filters = 64.

A batch size of 32 images is generally manageable on most modern GPUs, ensuring efficient use of hardware resources without causing memory issues. Additionally, this batch size introduces a moderate amount of noise in the gradient updates, providing some regularization effects without being too small. Empirical tests were conducted, and the training time was found to be acceptable, with the model converging effectively without overfitting. Additionally, the mean squared error (RMSE) was used as the loss function since the final layer of the network is a linear regression layer. Therefore, the RMSE tends to penalize larger errors more significantly, leading to better weight adjust-ments in the network compared to accuracy. The aver-

age RSME in the training stage was: 9.86.

This architecture was implemented on a PC with 32 GB of RAM, GPU NVIDIA Tesla K40c and a NVIDIA Tesla T4 working in parallel into an Ubuntu server environment.

PDM training parameters

- Training Images = 4800 images.
- Number of landmarks per example = 64.
- Explained variance = 0.9.

These parameters were selected as follows, the number of landmarks (64) are a good point density number and can represent the LV contour with good quality as depicted in Figure 6; normally the explained variance as proved in [9] is about 90 %, thus characterizing the most significant shape information provided by the PCA. Finally, the training set was the same as the one selected for the CNN training.
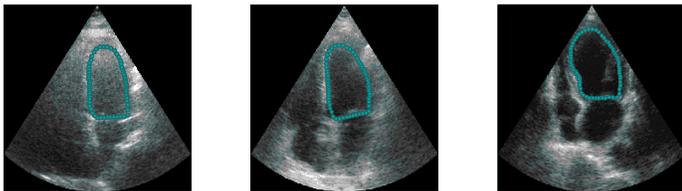

FIGURE 6. Example of landmarks sampled in the LV training images.

Segmentation results

In this section, we present the results obtained by our method on the test dataset. In Figures 7 and 8, are shown examples of image segmentation results for the CAMUS database during systole and diastole. Figures 9 and 10 illustrate the segmentation for both phases using images from the EchoNet Dynamic database. Additionally, Table 1 displays the average Dice coefficients for both test sets. Subsequently, Table 2 compares our method against others described in [20] and [22]. The average Dice value in this table represents the average obtained from the CNN on the CAMUS dataset

only. Furthermore, we decided to categorize the obtained segmentations based on their Dice coefficient into three categories: good, regular, and bad. The threshold used for this categorization is shown in Table 3. Meanwhile, in Figure 11a and 11b, the median for each category is observed for systole images from the CAMUS and EchoNet-Dynamic databases. Figures 11c and 11d display the median for each category for diastole images in the mentioned databases.
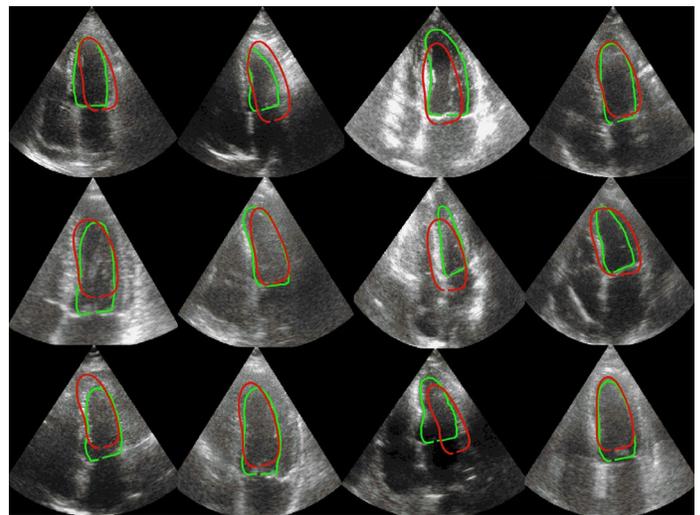

FIGURE 7. Systole CNN LV shape segmentations (red) vs. ground truth (green) for CAMUS database.
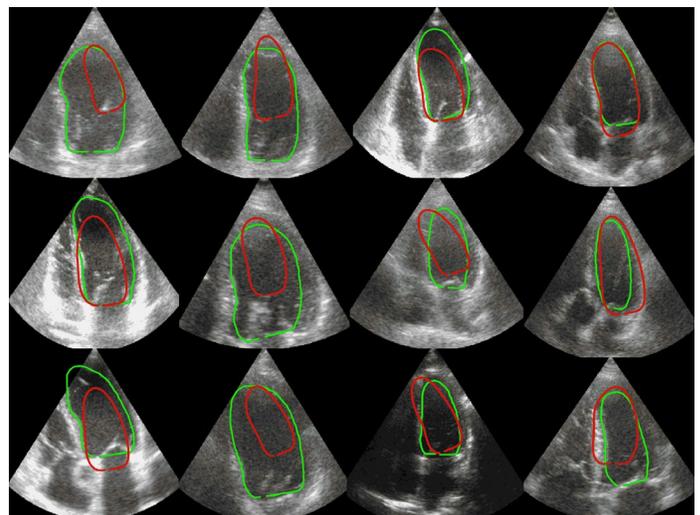

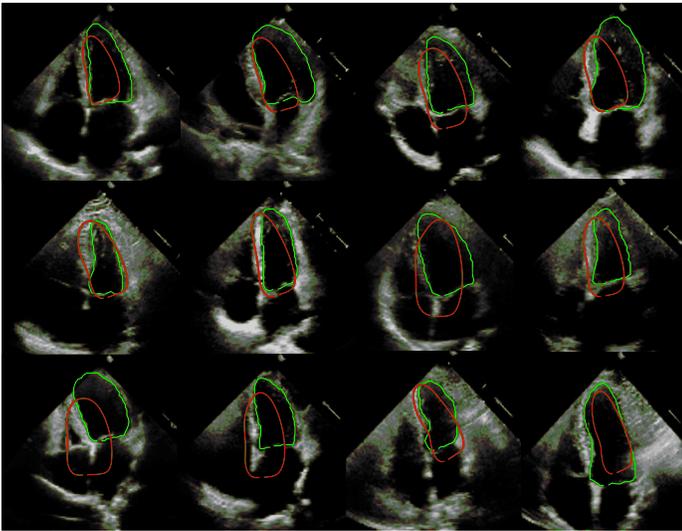**FIGURE 8.** Diastole CNN LV shape segmentations (red) vs. ground truth (green) for CAMUS database.

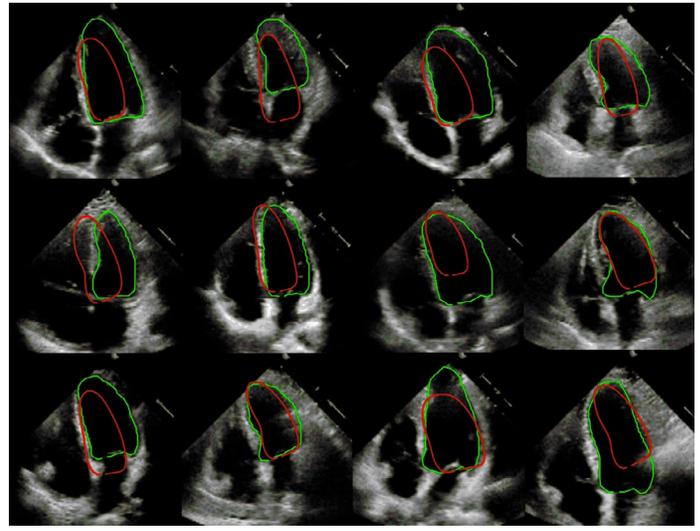**FIGURE 9.** Systole CNN LV shape segmentations (red) vs. ground truth (green) for EchoNet-dynamic database.

FIGURE 10. Diastole CNN LV shape segmentations (red) vs. ground truth (green) for EchoNet-dynamic database.

TABLE 1. Mean Haussdorff distance and mean Dice results for CAMUS and EchoNet-dynamic database

	CAMUS database		EchoNet-dynamic database	
Metric	Systole	Diastole	Systole	Diastole
Haussdorff distance (px)	25.92	37.55	28.95	31.20
Dice coefficient	0.71	0.66	0.64	0.65

TABLE 2. Comparison between the proposed U-Net versus other shape-based LV segmentation works

Method	Mean Dice coefficient systole	Mean Dice coefficient diastole
BEASM-SEMI [20]	0.86	0.92
BEASM-FULLY [20]	0.82	0.87
ASM [22]	0.82	0.81
FMR-ASM [22]	0.83	0.80
Proposed U-Net	**0.71**	**0.66**

TABLE 3. Category classification by Dice coefficient

Category	Threshold
Good	Dice coefficient $>= 0.80$
Regular	$0.70 <=$ Dice coefficient $<= 0.79$

Results analysis

In this section, we discuss the results presented in the previous section. The first point to address is the performance of our U-Net, both for the CAMUS database it was trained on and the EchoNet-dynamic database, which contains completely new images for the network

to segment. As observed in Figures 11a – 11d, the performance is very similar, which suggests that there is no significant overfitting due to data augmentation. We can also understand that the performance of the "good" category is competent in both cases (CAMUS and EchoNet-dynamic) with respect to the results pre-

sented in [20]. On the other hand, in Figures 7 to 10, it is evident that all segmentations produced by our method show shape characteristics very similar to those of the left ventricle (LV). They exhibit smoothed shapes located on the correct side of the image, with scale and rotation close to the expert annotation. Even in the case of poor segmentations (see Figure 12), the obtained contours preserve the shape qualities of the left ventricle. This is the effect achieved by the statistical shape constraint arising from training the network with parameters derived from a PDM of the LV. This represents an advantage over convolutional networks based on semantic segmentation, which, when they fail, produce classification errors like the one shown in Figure 1. Such misclassifications are challenging to correct since contour extraction yields noisy shapes

and sometimes contours are located far from the LV, a situation in which our method proves to be more robust. Another point to consider, which impacts our method, is the quality of the images. If ultrasound already poses a significant challenge due to speckle noise, the images acquired in both databases are not of the highest quality. This affects the segmentations obtained by our method and makes it difficult to find the optimal features corresponding to the pose and shape values during the network training hence the low Dice average in the tests performed. However, unlike semantic segmentation, these segmentations can be corrected by some other method that performs a LV contour fine-tuning.
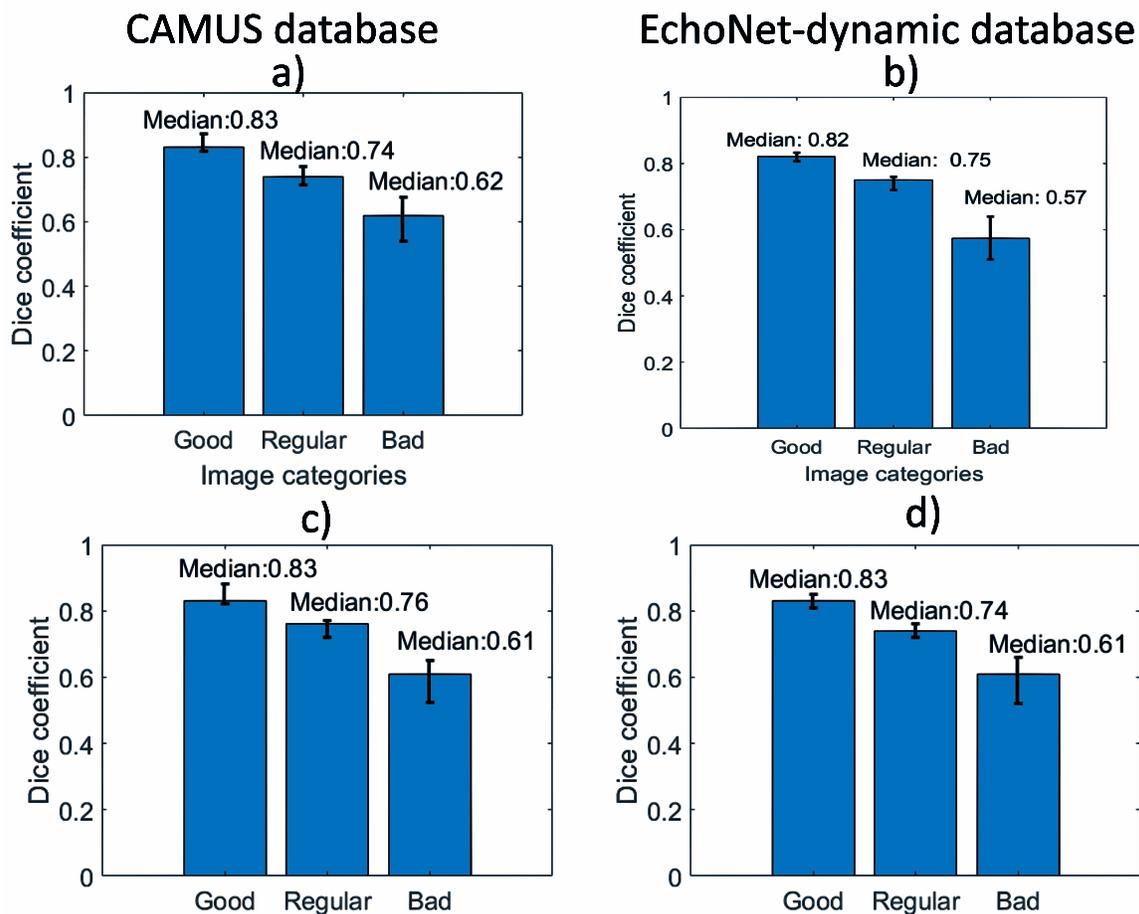

FIGURE 11. Median of each category for systole (11a, 11b) and diastole segmentations (11c, 11d) in CAMUS and EchoNet-dynamic database.
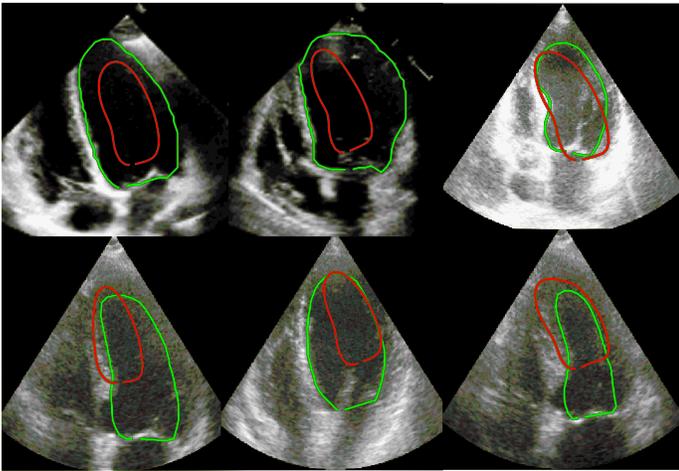
FIGURE 12. **Diastole CNN LV shape segmentations (red) vs. ground truth (green) for EchoNet-dynamic database.**

CONCLUSIONS

In this paper we presented an alternative way to the use of CNNs for LV shape and pose parameter prediction, taking advantage of convolution layers to find features extracted from the ultrasound image, resulting in segmentations with statistically valid shapes as the network is trained with the shape parameters of a PDM. Although, currently, the overall segmentation accuracy is not high, when high quality images are selected the accuracy of our method compares favorably to other previously published work. Our method avoids the appearance of blobs, and the fact that the segmentation is always a statistically valid shape allows it to be used as an initialization mechanism for another method to perform an adjustment and improve the segmentation starting from a smoothed shape and not from a noisy shape and in some cases far from the ventricle, as it would be the case with the extraction of the contour of a semantic segmentation CNN. In conclusion, this method explores the possibility of generating segmentations with statistical shape restrictions using the power of CNNs and can also be used as an automatic initialization method to later fine tune the predicted LV segmentation.

AUTHOR CONTRIBUTIONS

E.G.G. conceptualized the project, participated in the data curation, carried out formal analyses, contributed to investigation and the development of methodology and the use of specialized software, validated results and wrote the manuscript in their different stages of development. F.T.R. carried out formal analyses, participated in the development of methodology and in all the stages of the writing of the manuscript. J.L.P.G. carried out formal analyses, investigation, participated in the development of methodology and the writing, reviewing and editing of the manuscript. B.E.R. conceptualized and oversaw the project, carried out formal analyses, obtained resources for the project, and participated in the writing, reviewing and editing of the manuscript. F.A.C. conceptualized and oversaw the project, carried out formal analyses, obtained resources for the project, and participated in the writing, reviewing and editing of the manuscript. All Both authors reviewed and approved the final version of the manuscript.

REFERENCES

[1] N. Paragios and R. Deriche, "Geodesic active regions and level set methods for supervised texture segmentation," Int. J. Comput. Vis., vol. 46, no. 3, pp. 223-247, Feb. 2002, doi: **https://doi.org/10.1023/A:1014080923068**

[2] N. Paragios, "A level set approach for shape-driven segmentation and tracking of the left ventricle," IEEE Trans. Med. Imag., vol. 22, no. 6, pp. 773-776, Jun. 2003, doi: **https://doi.org/10.1109/TMI.2003.814785**

[3] J. C. Nascimento and J. S. Marques, "Robust shape tracking with multiple models in ultrasound images," IEEE Trans. Image Process., vol. 17, no. 3, pp. 392-406, Mar. 2008, doi: **https://doi.org/10.1109/TIP.2007.915552**

[4] V. Zagrodsky, V. Walimbe, C. R. Castro-Pareja, J. X. Qin, J.-M. Song, and R. Shekhar, "Registration-assisted segmentation of real-time 3-D echocardiographic data using deformable models," IEEE Trans. Med. Imaging., vol. 24, no. 9, pp. 1089-1099, Sep. 2005, doi: **https://doi.org/10.1109/tmi.2005.852057**

[5] B. Georgescu, X. S. Zhou, D. Comaniciu, and A. Gupta, "Database-guided segmentation of anatomical structures with complex appearance," in 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 2005, pp. 429-436 vol. 2, doi: **https://doi.org/10.1109/CVPR.2005.119**

[6] S. C. Mitchell, B. P. F. Lelieveldt, R. J. van der Geest, H. G. Bosch, J. H. C. Reiber, and M. Sonka, "Multistage hybrid active appearance model matching: Segmentation of left and right ventricles in cardiac MR images," IEEE Trans. Med. Imag., vol. 20, no. 5, pp. 415-423, May 2001, doi: **https://doi.org/10.1109/42.925294**

[7] O. Bernard, J. G. Bosch, B. Heyde, M. Alessandrini, et al., "Standardized Evaluation System for Left Ventricular Segmentation Algorithms in 3D Echocardiography," IEEE Trans. Med. Imaging, vol. 35, no. 4, pp. 967-977, Apr. 2016, doi: **https://doi.org/10.1109/TMI.2015.2503890**

[8] S. Leclerc, T. Grenier, F. Espinosa and O. Bernard, "A fully automatic and multi-structural segmentation of the left ventricle and the myocardium on highly heterogeneous 2D echocardiographic data," 2017 IEEE International Ultrasonics Symposium (IUS), Washington, DC, USA, 2017, pp. 1-4, doi: **https://doi.org/10.1109/ULTSYM.2017.8092797**

[9] T. F. Cootes, C. J. Taylor, D. H. Cooper, J. Graham, "Active shape models their training and application," Comput. Vis. Image Underst., vol. 61, no. 1, pp. 38-59, Jan. 1995, doi: **https://doi.org/10.1006/cviu.1995.1004**

[10] Y. Ali, S. Beheshti, F. Janabi-Sharifi, "Echocardiogram segmentation using active shape model and mean squared eigenvalue error," Biomed. Signal Process. Control, vol. 69, art. no. 102807, Aug. 2021, doi: **https://doi.org/10.1016/j.bspc.2021.102807**

[11] M. A. Shoaib, J. Huang Chuah, R. Ali, K. Hasikin, et al., "An Overview of Deep Learning Methods for Left Ventricle Segmentation," Comput. Intell. Neurosci., vol. 2023, art. no. 4208231, 2023, doi: **https://doi.org/10.1155/2023/4208231**

[12] N. Savioli, M. S. Vieira, P. Lamata, G. Montana, "Automated Segmentation on the Entire Cardiac Cycle Using a Deep Learning Work - Flow," in 2018 Fifth International Conference on Social Networks Analysis, Management and Security (SNAMS), Valencia, Spain, 2018, pp. 153-158, doi: **https://doi.org/10.1109/SNAMS.2018.8554962**

[13] X. Zou, Q. Wang, T. Luo, "A novel approach for left ventricle segmentation in tagged MRI," Comput. Electr. Eng., vol. 95, art. no. 107416, Oct. 2021, doi: **https://doi.org/10.1016/j.compeleceng.2021.107416**

[14] T. Wech, M. J. Ankenbrand, T. A. Bley, J. F. Heidenreich, "A data-driven semantic segmentation model for direct cardiac functional analysis based on undersampled radial MR cine series," Magn. Reason. Med., vol. 87, no. 2, pp. 972-983, Feb. 2022, doi: **https://doi.org/10.1002/mrm.29017**

[15] G. Veni, M. Moradi, H. Bulu, G. Narayan, and T. Syeda-Mahmood, "Echocardiography segmentation based on a shape-guided deformable model driven by a fully convolutional network prior," 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA, 2018, pp. 898-902, doi: **https://doi.org/10.1109/ISBI.2018.8363716**

[16] W.-Y. Hsu, "Automatic Left Ventricle Recognition, Segmentation and Tracking in Cardiac Ultrasound Image Sequences," IEEE Access, vol. 7, pp. 140524-140533, 2019, doi: **https://doi.org/10.1109/ACCESS.2019.2920957**

[17] Y. Li, W. Lu, P. Monkam, Z. Zhu, W. Wu, M. Liu, "LVSnake: Accurate and robust left ventricle contour localization for myocardial infarction detection," Biomed. Signal Process. Control, vol. 85, art. no. 105076, Aug. 2023, doi: **https://doi.org/10.1016/j.bspc.2023.105076**

[18] S. Peng, W. Jiang, H. Pi, X. Li, H. Bao, and X. Zhou, "Deep Snake for Real-Time Instance Segmentation," 2020, arXiv: 2001.01629, doi:**https://doi.org/10.48550/arXiv.2001.01629**

[19] O. Ronneberger, P. Fischer, T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Munich, Germany, 2015, pp. 234-241, doi: **https://doi.org/10.1007/978-3-319-24574-4_28**

[20] S. Leclerc, E. Smistad, J. Pedrosa, A. Østvik, et al., "Deep Learning for Segmentation using an Open Large-Scale Dataset in 2D Echocardiography," IEEE Trans. Med. Imaging, vol. 38, no. 9, pp. 2198-2210, Sep. 2019, doi: **https://doi.org/10.1109/TMI.2019.2900516**

[21] EchoNet-Dynamic: a Large New Cardiac Motion Video Data Resource for Medical Machine Learning, github, 2019. [Online]. Available: **https://echonet.github.io/dynamic/index.html#motivation**

[22] E. Galicia Gomez, F. Torres Robles, B. Escalante Rmirez, J. Olveres, F. Arámbula Cosío, "Full multi resolution active shape model for left ventricle segmentation," in 17th International Symposium on Medical Information Processing and Analysis, Campinas, Brazil, 2021, doi: **https://doi.org/10.1117/12.2606252**

[23] M. B. Stegmann, D. Delgado Gomez. (2002). A Brief Introduction to Statistical Shape Analysis [Online]. Available: **http://www2.imm.dtu.dk/pubdb/edoc/imm403.pdf**

[24] A. Cervantes-Guzmán, K. McPherson, J. Olveres, C. F. Moreno-García, F. Torres Robles, E. Elyan, B. Escalante-Ramírez, "Robust cardiac segmentation corrected with heuristics," PLoS One, vol. 18, no. 10, art. no. E0293560, Oct. 2023, doi: **https://doi.org/10.1371/journal.pone.0293560**