Article

Not peer-reviewed version

CLARISA: Connexin-43 Lateralization Automated ROI-Based Image Signal Analyzer

Daniel Gattari , Joseba Sancho-Zamora , Debora Chan , Emiliano Diez , Mariano Llamedo Soria , Mario Rossi *

Posted Date: 2 May 2026

doi: 10.20944/preprints202605.0043.v1

Keywords: connexin-43; lateralization; deep learning; multi-scale classification; fluorescence microscopy; automatic quantification

 doi:10.20944/preprints202605.0043.v1

Article

CLARISA: Connexin-43 Lateralization Automated ROI-Based Image Signal Analyzer

Daniel Gattari [1], Joseba Sancho-Zamora [2], Debora Chan [1], Emiliano Diez [3], Mariano Llamedo Soria [4] and Mario Rossi [1,5,*]

[1] Faculty of Engineering, Austral University, Pilar, Buenos Aires, Argentina
[2] Tecnun School of Engineering, Universidad de Navarra, Donostia, Spain
[3] Institute of Experimental Medicine and Biology of Cuyo (IMBECU), CONICET, Mendoza, Argentina
[4] Electronics Department, National Technological University, Buenos Aires, Argentina
[5] Functional Genomics and Data Science, Institute for Translational Medicine Research (IIMT), CONICET-Universidad Austral, Pilar, Buenos Aires, Argentina
* Correspondence: mrossi1@austral.edu.ar

Abstract

Connexin-43 (CX43) lateralization in ventricular myocardium has been associated with abnormal impulse propagation and increased arrhythmia susceptibility. Its quantitative assessment in histological sections remains challenging because of the difficulty of segmenting individual cardiomyocytes and the reliance of previous methods on geometric rules applied to segmented cell profiles. Here, we present CLARISA, a deep learning framework for classifying CX43-positive regions as either terminal or lateralized directly from fluorescence images, without requiring cardiomyocyte segmentation. An expert-annotated dataset was generated from left-ventricular cryosections of Wistar rat hearts, in which CX43-positive regions were labeled according to their distribution pattern. A dual-stream convolutional classifier based on EfficientNetV2-S was trained to capture both the local and contextual morphology of each region. In addition, an inference module applicable to whole tissue sections was developed to generate spatial lateralization probability maps and global percent lateralization estimates consistent with expert annotation. On the test set, CLARISA achieved a ROC-AUC of 0.905 and a PR-AUC of 0.810. These results support the feasibility of automated assessment of CX43 distribution patterns without explicit cardiomyocyte segmentation. The complete codebase is publicly available, together with access to the pretrained model and the image data used in this study. The Hugging Face model card reports the same held-out test metrics and states that the checkpoint is intended to be used with the main repository.

Keywords: connexin-43; lateralization; deep learning; multi-scale classification; fluorescence microscopy; automatic quantification

1. Introduction

Cardiac impulse propagation depends on multiple structural and functional determinants. Among these, cardiomyocyte size and shape play a central role [1,2]. At the cellular level, membrane area contributes to capacitive properties, whereas cell volume influences intracellular resistance. Beyond the single cell, conduction is further modulated by tissue-level organization, including the spatial arrangement of myocytes, non-myocyte cells, and connective tissue within the extracellular space [3]. At the molecular level, impulse propagation is critically shaped by the expression and spatial distribution of connexins.

Connexins form intercellular channels at gap junctions, providing direct electrical and metabolic coupling between adjacent cardiomyocytes and thereby sustaining the syncytial behavior of the myocardium [4]. In healthy ventricular tissue, connexin-43 (CX43) is predominantly concentrated at the intercalated discs located at the longitudinal ends of cardiomyocytes. Under pathological or stress

 doi:10.20944/preprints202605.0043.v1

conditions, however, CX43 may redistribute toward the lateral cell borders, a phenomenon commonly referred to as lateralization [5].

Because the spatial arrangement of gap junctions directly influences both the strength and anisotropy of intercellular coupling, changes in CX43 distribution are considered functionally relevant. In diseased myocardium, increased lateralization, often accompanied by an overall reduction in CX43 expression, has been associated with abnormal impulse propagation and increased susceptibility to arrhythmias [6]. However, the functional consequences of this redistribution remain debated, as some studies suggest that laterally displaced CX43 may partially compensate for structural inhomogeneities by preserving some degree of coupling in remodeled regions [7]. This unresolved issue highlights the need for quantitative approaches capable of characterizing CX43 distribution across large tissue areas.

Quantitative assessment of CX43 distribution in histological sections poses substantial practical challenges. The number of intercellular junctions within the myocardium far exceeds what can be comprehensively evaluated by conventional microscopy, while manual annotation is labor-intensive and therefore usually restricts analysis to small regions of interest. As a result, many studies may fail to capture the spatial heterogeneity of connexin remodeling across an entire tissue section [8]. A further challenge lies in the segmentation of individual cardiomyocytes in intact tissue. Their complex branching morphology, dense packing, and irregular boundaries make reliable delineation difficult even for trained observers [9]. In standard histological preparations, oblique sectioning and limited membrane contrast further increase the uncertainty associated with identifying individual cellular profiles [10].

These limitations have motivated the development of automated and semi-automated image analysis methods. In cardiac tissue, deep learning approaches have been applied to histopathological classification and cellular quantification [11–13]. More specialized methods have also been proposed to assess gap junction distribution in rat ventricular myocardium [14] and to evaluate CX43 localization through co-localization with N-cadherin in atrial tissue [15]. Automated quantification of intercellular coupling has likewise been reported in both non-cardiac cell cultures [16] and cardiac co-culture systems [17]. Despite these advances, no currently available method provides an automated framework for classifying CX43 distribution patterns across large myocardial tissue sections, particularly for distinguishing terminal from lateralized signal.

A method specifically designed for the automated quantification of CX43 lateralization in fluorescence immunohistochemistry images of ventricular myocardium has previously been reported [18]. For clarity, this method is hereafter referred to as MARTA. In that approach, explicit segmentation of individual cardiomyocytes was required as a prerequisite for quantification. Once cell contours had been identified, each cell area was subdivided into terminal and central compartments using geometric rules, and CX43 signal was assigned accordingly to estimate a lateral-to-total ratio. This design imposed several limitations. First, quantification was restricted to regions in which cardiomyocytes could be successfully segmented, leaving any CX43 signal outside delineated cells unassigned. Second, the method relied on a fixed geometric partition of the cell profile, assuming that compartment boundaries derived from a bounding rectangle adequately represented the underlying cellular morphology. Third, the pipeline consisted of a predefined sequence of morphological operations, including dilation, erosion, and contour extraction, and therefore did not learn from data, potentially limiting its robustness across images with different staining characteristics or tissue properties. Finally, extending this framework to supervised learning would require instance-level segmentation masks of individual cardiomyocytes, a type of annotation that is particularly difficult to generate in intact myocardial tissue.

In the present work, we introduce CLARISA (Connexin Lateralization Automated ROI-based Image Signal Analyzer), an alternative framework designed to address these limitations. Rather than segmenting cardiomyocytes and assigning CX43 signal to predefined geometric compartments, CLARISA treats CX43-positive regions as the primary unit of analysis and classifies them directly as terminal or lateralized according to their appearance within the surrounding tissue context. This

 doi:10.20944/preprints202605.0043.v1

reformulation substantially simplifies the annotation process: instead of requiring cell-level segmentation masks, it relies only on expert identification of CX43-positive regions and assignment of each region to one of two classes. As a result, it enables the construction of a labeled dataset suitable for supervised learning.

The availability of such data makes it possible to train a deep learning classifier that learns the morphological features distinguishing lateralized from terminal CX43 signal, rather than relying on a fixed sequence of image-processing operations. At inference, the method classifies all automatically detected CX43-positive regions within a tissue section, independently of whether they fall within a successfully segmented cell boundary.

The proposed framework comprises three main components: (i) generation of an expert-annotated dataset of CX43-positive regions labeled as terminal or lateralized; (ii) training of a deep learning classifier, based on a pretrained backbone, to predict the lateralization probability of individual CX43-positive regions from their local and contextual image appearance; and (iii) an inference module for application to full tissue sections, in which CX43-positive regions are automatically detected and individually classified, enabling computation of the proportion of lateralized signal relative to the total CX43-positive area. In addition, spatial probability maps are generated to facilitate visual inspection of predicted lateralization patterns across the tissue. The entire pipeline operates without explicit cardiomyocyte segmentation.

The aim of this study was to develop and evaluate a methodological framework for the automated classification of CX43 distribution patterns in fluorescence immunohistochemistry images of rodent ventricular myocardium. The approach was validated at the level of individual CX43-positive regions using held-out tissue section sets, and its applicability to whole tissue sections was demonstrated through the inference module described above, including comparison at the inference level with the previously developed MARTA framework [18]. The method, together with the annotated dataset and the code required to reproduce the pipeline, was made publicly available through a GitHub repository, with the aim of providing a baseline that can be adapted, extended, or retrained for different tissue preparations, species, or staining protocols.

2. Materials and Methods

2.1. Data Acquisition and Histological Imaging

Hearts from 3-month-old male Wistar rats were excised and snap-frozen in liquid nitrogen as previously described [5]. For connexin-43 (CX43) immunodetection, 10-µm cryosections of the left ventricle were washed in phosphate-buffered saline (PBS), fixed in ice-cold methanol, permeabilized with 0.3% Triton X-100 in PBS, and blocked with 1% bovine serum albumin in PBS. Sections were incubated with a primary mouse anti-CX43 antibody (MAB3068, Chemicon International, Temecula, CA, USA) and stained for F-actin (ab112124, Abcam, Cambridge, UK), followed by incubation with a goat anti-mouse FITC-conjugated secondary antibody (111-095-003, Jackson ImmunoResearch, West Grove, PA, USA). Slides were mounted with Vectashield mounting medium (H-1200, Vector Laboratories, Burlingame, CA, USA) and imaged using a Zeiss Axio Imager Z2 microscope equipped with ApoTome.2 (Carl Zeiss, Jena, Germany).

Images used for classifier training, validation, and testing were acquired at a spatial resolution of 0.227 µm/pixel. An additional tissue section, used exclusively for the evaluation described in Section 2.7, was obtained from the same cohort under the same experimental protocol but imaged at a spatial resolution of 0.3 µm/pixel.

2.2. Expert Annotation and Dataset Generation of CX43-Positive Regions

The annotated dataset was constructed from two global fluorescence images, from which six tissue sections were selected for annotation (Figure 1). Each global image was subdivided into a regular 4×3 grid of non-overlapping tissue sections. The six sections selected for annotation occupied different positions within this grid, as indicated in Table 1.

 doi:10.20944/preprints202605.0043.v1

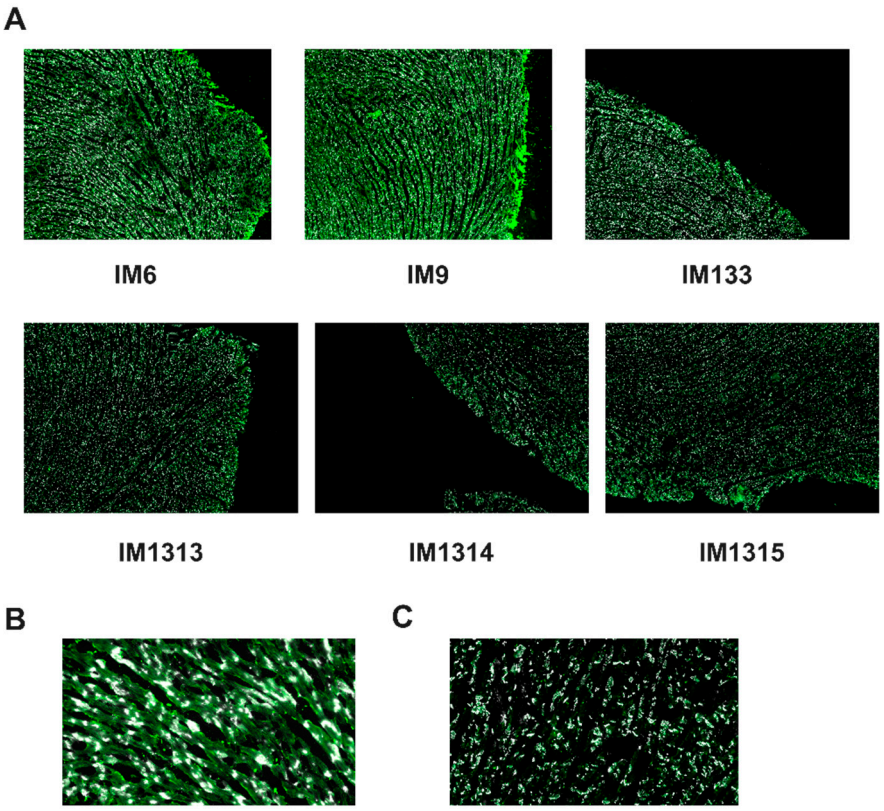

Figure 1. Tissue sections used for expert annotation, with representative close-up views of CX43 staining. (A) Overview of the six histological sections from which all annotated CX43-positive regions of interest were drawn. The dataset was partitioned at the section level into training (IM6, IM9, IM133), validation (IM1315), and held-out test (IM1313, IM1314) subsets, as described in Section 3.5.1. All sections were obtained from the left ventricle of 3-month-old male Wistar rats and imaged under the conditions detailed in Section 3.1; green signal corresponds to F-actin and white signal to CX43, and within the myocardium the black background corresponds to interstitial space. (B, C) Illustrative magnified views showing the appearance of individual CX43-positive regions at higher zoom, extracted from IM6 (B) and IM1313 (C). These magnifications are provided for visualization only and were not used as additional ROIs at any stage of model development.

All annotations were performed by a single domain expert according to predefined morphological criteria (Figure 2, Panel A). CX43-positive regions were classified as terminal when the signal displayed a morphology and spatial distribution consistent with localization at the longitudinal ends of cardiomyocytes, and as lateralized when the signal exhibited a pattern redistributed along the lateral borders of the cellular profile.

For each annotation, the expert visually identified a CX43-positive region, assigned the corresponding class label, and manually selected a window around it. This window was subsequently processed by a computational pipeline to delineate the CX43-positive area and derive standardized bounding-box coordinates. To this end, the selected window was first converted to grayscale and thresholded at a fixed intensity of 180 on the 8-bit scale to isolate the brightest pixels corresponding to the fluorescent signal. Morphological opening was applied to remove small spurious components, followed by dilation to consolidate the remaining signal area. A bounding rectangle was then fitted around the resulting closed region and expanded by 40 pixels (approximately 9 μm at the imaging resolution reported in Section 2.1) in all directions to reduce the risk of excluding peripheral signal.

In this study, a region of interest (ROI) was defined as a rectangular image patch centered on a CX43-positive region and represented by bounding-box coordinates of the form (x_1, y_1, x_2, y_2),

 doi:10.20944/preprints202605.0043.v1

together with an associated class label. The final dataset comprised 3139 ROIs, of which 1969 were classified as terminal and 1170 as lateralized.

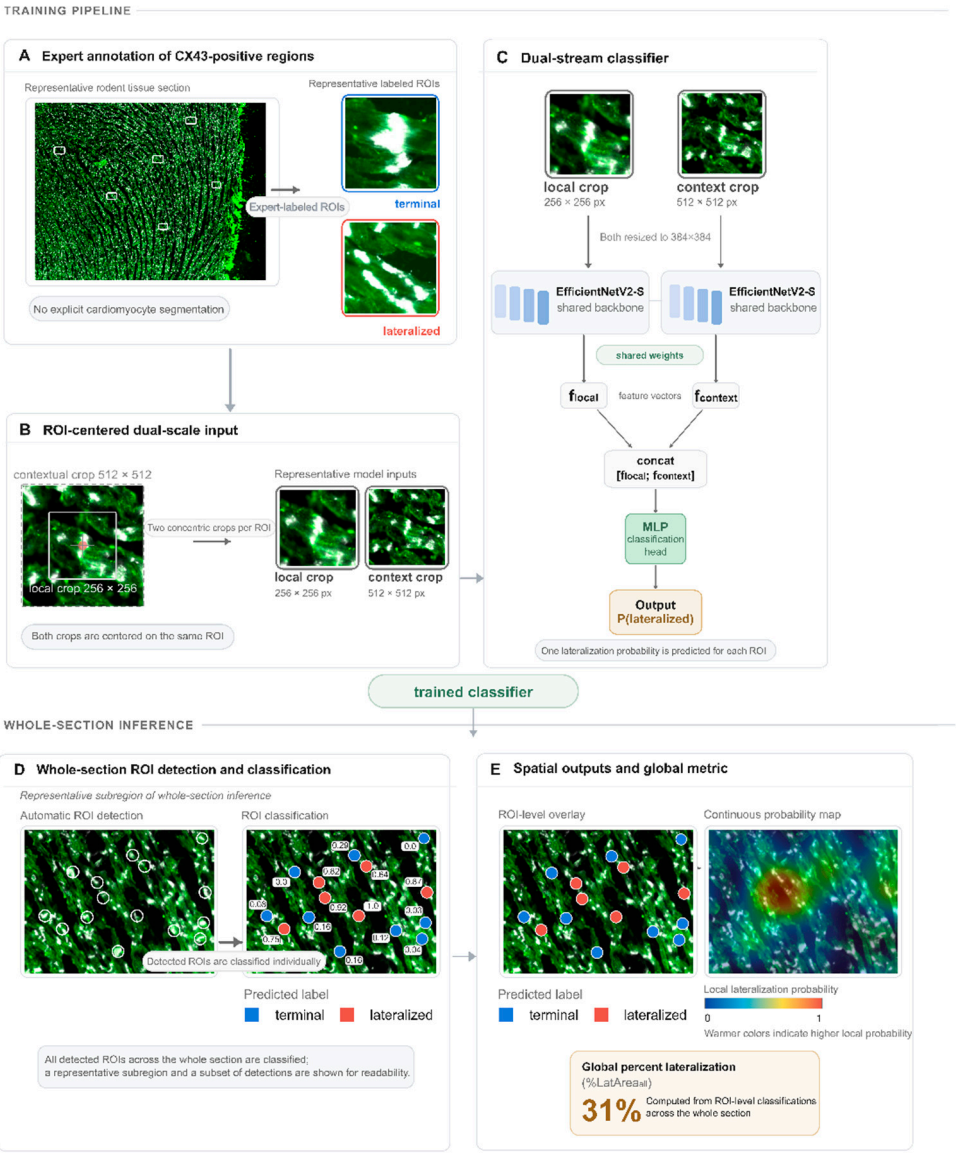

Figure 2. Overview of the proposed methodological framework CLARISA for automated quantification of CX43 lateralization in histological images of cardiac tissue. The workflow is organized in two stages: a training pipeline (top) in which a classifier is learned from expert-annotated data, and a whole-section inference pipeline (bottom) in which the trained classifier is applied to previously unseen tissue sections. (A) Expert annotation of CX43-positive regions. On representative rodent ventricular tissue sections, the expert manually selected regions of interest (ROIs) around CX43-positive signal and assigned each of them to one of two classes: terminal, when the signal was located at the longitudinal ends of the cardiomyocyte, or lateralized, when it was located along its lateral membrane. Annotation was performed directly on the fluorescence image, without explicit segmentation of cardiomyocyte contours. (B) ROI-centered dual-scale input. For each annotated ROI, two spatially aligned square crops centered on the same location were extracted from the tissue section: a local crop capturing the CX43-positive region itself and its immediate surroundings, and a context crop capturing a larger portion of surrounding tissue. (C) Dual-stream classifier. The local and context crops, both resized to 384 × 384 px, were encoded by an EfficientNetV2-S backbone with shared weights across the two streams, yielding two feature vectors that were concatenated and passed to a multi-layer perceptron (MLP) classification head. The classifier output was a single scalar, P(lateralized), representing the predicted probability that the input ROI belongs to

 doi:10.20944/preprints202605.0043.v1

the lateralized class. (D) Whole-section ROI detection and classification. When applied to a new tissue section, the trained classifier is embedded in an automated pipeline that first detects candidate CX43-positive regions by morphological processing of the fluorescence image, without any prior expert intervention, and then classifies each detected ROI individually, producing a predicted probability and a final class label per region. The panel illustrates this step on a representative subregion of the whole section, with per-ROI probabilities shown on a subset of detections for readability. (E) Spatial outputs and global metric. The per-ROI classifications are aggregated into two complementary representations of the predicted lateralization pattern across the tissue section — a discrete ROI-level overlay coloring each detected region by its assigned class, and a continuous probability map obtained by Gaussian spatial interpolation of the per-ROI probabilities (warmer colors indicate higher local lateralization probability) — and summarized by a single global metric, $\%LatArea_{all}$, defined as the fraction of the total detected CX43-positive area assigned to the lateralized class. The illustrative value shown in the panel corresponds to the representative subregion and does not reflect any specific experimental result reported in the paper.

Table 1. Distribution of annotated CX43-positive regions across global images and tissue sections.

Global image	Tissue section	Grid Position	Terminal CX43-positive regions	Lateralized CX43-positive regions
Global 1	IM9	row 2, col 3	1161 (77.19%)	343 (22.81%)
Global 1	IM6	row 3, col 3	567 (42.06%)	781 (57.94%)
Global 2	IM133	row 1, col 3	110 (96.49%)	4 (3.51%)
Global 2	IM1313	row 3, col 3	43 (89.58%)	5 (10.42%)
Global 2	IM1314	row 4, col 1	38 (66.67%)	19 (33.33%)
Global 2	IM1315	row 4, col 2	50 (73.53%)	18 (26.47%)

2.3. Centered Cropping Extraction and Dual-Scale Input Generation

Each annotated bounding box was transformed into image crops suitable as input to the classifier. The crop center was defined as the integer midpoint of the bounding-box coordinates. Two spatially aligned crops were then extracted around this center: a local crop of 256 × 256 pixels capturing the immediate CX43 neighborhood, and a contextual crop of 512 × 512 pixels providing the surrounding histological context (Figure 2, Panel B). At the imaging resolution reported in Section 2.1, these crop sizes enabled simultaneous representation of two complementary scales: the smaller crop preserved sufficient detail to assess the local morphology of the CX43-positive region and its position within the myocardial fiber, whereas the larger crop provided contextual information on CX43 expression in neighboring fibers. following previous multi-scale designs proposed for medical image analysis [19,20].

Together, these two centered views constituted the dual-scale input representation used in the classification framework. Both crops were centered on the same coordinates, ensuring spatial correspondence across scales and consistent positioning of the CX43-positive region within the input, in a manner consistent with the downstream inference pipeline.

2.4. Preprocessing and Data Augmentation

All image crops were resized to 384 × 384 pixels and normalized using the ImageNet channel-wise mean and standard deviation [21]. During training, standard geometric data augmentations were applied, including horizontal and vertical flipping, affine translation, isotropic scaling and limited rotation, following general recommendations for image augmentation in deep-learning pipelines [22].

For the dual-scale input setting, identical stochastic transformations were applied to both synchronized crops to preserve spatial alignment between the local and contextual views. No augmentations were applied during validation or test evaluation.

 doi:10.20944/preprints202605.0043.v1

2.5. CX43 Lateralization Classifier

2.5.1. Data Splitting Strategy

The annotated dataset was partitioned at the slide level, such that all ROIs from a given tissue section were assigned to the same subset. This design prevented ROIs from the same section — which share staining conditions, tissue orientation, and background characteristics — from appearing in both training and evaluation subsets, thereby reducing the risk of data leakage. Given the limited number of available sections, a fixed split was adopted: IM6, IM9, and IM133 were assigned to the training set, IM1315 to the validation set; and IM1313 together with IM1314 to the held-out test set. The corresponding ROI counts and class distributions are summarized in Table 2.

Although IM133 originated from the same global image as IM1313, IM1314, and IM1315, it occupied a non-adjacent position within the tissue grid and displayed distinct local tissue characteristics (Table 1). It was therefore considered informative for increasing morphological and staining variability during training, while preserving separation at the section level for validation and testing.

This partitioning strategy resulted in relatively small validation and test sets, and performance estimates should therefore be interpreted with appropriate caution. Nevertheless, slide-level splitting was preferred over ROI-level splitting because it provides a more conservative assessment of generalization. For completeness, ROI-level partitioning was also evaluated, and the corresponding results are reported in Supplementary Section S4.

Table 2. Dataset partitioning at the slide level

Subset	Sections	n (ROIs)	Class terminal (%)	Class lateralized (%)
Training	IM6, IM9, IM133	2966	61.97	38.03
Validation	IM1315	68	73.53	26.47
Test	IM1313, IM1314	105	77.14	22.86

2.5.2. Model Architecture

The classifier (Figure 2C) operates on the preprocessed dual-scale input described in Sections 2.3 and 2.4, and outputs a lateralization probability, P(lateralized), through a shared feature extractor followed by a joint classification head.

Feature extraction is performed with an EfficientNetV2-S backbone pretrained on ImageNet [23]. The original classification layer is removed, and global average pooling is applied to the final feature map, yielding a 1280-dimensional feature vector for each crop. In the dual-stream configuration, the local and contextual crops are passed sequentially through the same backbone with shared weights, producing two feature vectors f_{local} and $f_{context}$. This shared-weight design ensures that both spatial scales are encoded within a common feature space while reducing the number of trainable parameters relative to architectures with independent backbones.

The two feature vectors are then concatenated to form a fused representation $f = [f_{local}; f_{context}]$, of dimension 2560, which is fed into the classification head to generate the final prediction (see Figure 2C). Because the prediction is derived from this joint representation, optimization updates the shared backbone using information from both input scales, encouraging the extraction of features that are informative at both the local and contextual levels. Alternative input formulations (local-only and contextual-only) and fusion strategies were also evaluated during architecture optimization (Section 2.5.4).

2.5.3. Training Strategy and Staged Fine-Tuning

Model training was performed using binary cross-entropy with logits as the loss function. To account for class imbalance, the loss was weighted by the inverse class frequency computed on the training set, with the positive-class weight defined as $w^+ = n_0 / n_1$, where n_0 and n_1 denote the number

 doi:10.20944/preprints202605.0043.v1

of terminal and lateralized samples in the training subset, respectively [24]. This weight was further scaled by a factor of 1.1 to slightly increase the contribution of the minority class during optimization.

Training followed a staged fine-tuning strategy to progressively adapt the pretrained backbone to task-specific morphological features while mitigating catastrophic forgetting [25,26]. In stage 1, the backbone was frozen and only the classification head was optimized. In stage 2, the last two backbone blocks were unfrozen to allow partial adaptation of high-level features. In stage 3, the entire backbone was unfrozen, and the full network was fine-tuned jointly.

Optimization was performed using AdamW with a weight decay of 1.0 × 10−4. A batch size of 16 was used. Training was conducted for 6, 4 and 10 epochs in stages 1, 2 and 3, respectively. Different learning rates were assigned to different parts of the model to allow faster adaptation of the newly initialized classification head while preserving the more general pretrained representations in earlier backbone layers. The classification head was trained with a learning rate of 9.83×10^{-4}, the last unfrozen backbone block with 3.44×10^{-4}, and the remaining backbone layers with 1.15×10^{-4}.

2.5.4. Architecture and Hyperparameter Optimization

Architecture and hyperparameter optimization was performed using Optuna [27] with a Tree-structured Parzen Estimator (TPE) sampler [28]. Optimization proceeded in two sequential stages: an architecture search followed by a fine-tuning search focused on training-related hyperparameters. A detailed description of the optimization framework, including fixed parameters and search spaces, is provided in Supplementary Section S1.1.

In the first stage, alternative model formulations were evaluated while keeping the training configuration fixed. The explored variables included the input representation, the fusion strategy for dual-scale models, the classifier-head type, and the dimensionality and dropout of the classification head. In the second stage, the selected architecture was fixed, and optimization focused on training-related parameters, including dropout, the number of unfrozen backbone blocks during stage 2, the allocation of epochs across the three training stages, the learning-rate hierarchy across model components, and weight decay. The corresponding fixed parameters and search spaces are reported in Supplementary Tables 1 and 2, and the full trial-wise results are provided in Supplementary File 1.

Each trial consisted of training a candidate model using the staged fine-tuning procedure described in Section 2.5.3 while monitoring performance on the validation set. The primary optimization objective was the minimum validation loss achieved during training, and validation AUC was recorded as a complementary metric. Thus, candidate configurations were compared according to their best validation-loss checkpoint rather than according to final-epoch performance.

All trials used the same train–validation split to ensure comparability across configurations. Given the limited dataset size, this procedure should be interpreted as a controlled model selection process rather than an exhaustive exploration of the search space. The best-performing configuration was selected on the basis of minimum validation loss, retrained using the same settings, and the checkpoint with the lowest validation loss in the final run was retained for test evaluation. Further details are provided in Supplementary Section S1.1.3.

2.5.5. Architecture and Hyperparameter Optimization

After model selection on the training and validation sets, the final model was evaluated on the held-out test set, which was not used during model development. For each annotated ROI, the model produced a probability of lateralization.

Final classifier performance on the test set was assessed using both threshold-independent and threshold-dependent metrics. Threshold-independent performance was quantified by the area under the receiver operating characteristic curve (ROC-AUC) and the area under the precision-recall curve (PR-AUC, computed as average precision); the latter has been recommended as more informative than ROC-AUC for evaluating binary classifiers on imbalanced datasets [29]. Threshold-dependent performance was evaluated at the default decision threshold of 0.5, assigning ROIs to the lateralized

 doi:10.20944/preprints202605.0043.v1

class when P(lateralized) ≥ 0.5 and to the terminal class otherwise, and reporting precision and recall for both classes.

2.6. Inference Module for Whole Images

2.6.1. Automatic Detection of CX43-Positive Regions

The inference pipeline operates on whole tissue sections and therefore includes an initial detection step to identify candidate CX43-positive regions for subsequent classification (Figure 2D). This step used the same morphology-based detection pipeline described for annotation in Section 2.2, but applied exhaustively to the full image rather than to manually selected windows. The image was first converted to grayscale, thresholded at a fixed intensity value, and processed by morphological opening followed by dilation. External contours were then extracted from the resulting binary mask. For each detected contour, a bounding rectangle was computed, isotropically expanded by a fixed margin, and clipped to the image boundaries. The coordinates of all retained bounding boxes were stored for downstream processing.

The detector exposes three adjustable parameters that govern its behavior: the intensity threshold used for binarization, and the sizes of the structuring elements used for the opening and dilation operations. Their default values (Table 3) reproduce the configuration used during dataset generation (Section 2.2) and were selected for the staining and imaging conditions described in Section 2.1. Because these operations are directly linked to the visual appearance of the CX43 signal, the default values may require adjustment when CLARISA is applied to tissue preparations with different staining intensity, background levels, or CX43 expression patterns. Practical guidance for parameter adjustment is provided in Table 3. A fourth parameter, the bounding-box expansion margin, was kept fixed at 40 pixels throughout this work; it adds a small contextual border around each detected contour and typically does not require modification.

Table 3. Adjustable parameters of the CX43-positive region detector and practical guidance for tuning.

Parameter	Role	Default	Practical guidance for adjustment
thresh_value	Intensity threshold (0–255) used to binarize the grayscale image. Pixels above this value are considered CX43-positive.	180	Decrease (e.g. 100–150) if staining is weak or the image is dark; increase (e.g. ≥200) if staining is strong or background noise is high.
kernel_open	Size (px) of the square structuring element used for morphological opening. Removes isolated bright regions smaller than the kernel.	9	The opening operation is destructive: any bright region removed at this step cannot be recovered by subsequent operations. When CX43 expression is low, small detections may correspond to genuine but faint CX43 signal rather than to noise; a large *kernel_open* therefore risks eliminating relevant information irreversibly. Decrease when expression is low or fragmented to preserve sensitive signal; increase when staining is strong and small detections are expected to be background noise. Set to 1 to disable.
kernel_dilate	Size (px) of the square structuring	5	When CX43 expression appears fragmented or sparse — either because signal intensity is weak or because

 doi:10.20944/preprints202605.0043.v1

	element used for morphological dilation. Expands detected regions and can merge nearby detections into a single ROI.	positive regions are spatially scattered — the raw detection may split a single biologically coherent region into multiple small ROIs separated by dark gaps. Increasing *kernel_dilate* allows these fragments to grow and reconnect into ROIs that represent the underlying biological region as a whole. Decrease if distinct cellular ROIs are being merged incorrectly. Set to 1 to disable.

2.6.2. Classification of Detected Regions

Each automatically detected ROI was converted into classifier input using the same centered-crop and preprocessing procedure described for training (Sections 2.3 and 2.4). When CLARISA is applied to images acquired at a spatial resolution different from that of the development set (Section 2.1), the default crop sizes (256 × 256 px for the local crop and 512 × 512 px for the context crop) must be rescaled according to the ratio between the default and target resolutions, in μm/pixel, so that the physical field of view represented by each crop is preserved. No other aspect of the classification procedure depends on image resolution.

During inference, the trained classifier produced a lateralization probability P(lateralized) ∈ [0, 1] for each detected ROI (Figure 2D), from which the terminal probability was obtained as P(terminal) = 1 − P(lateralized). A final class label was assigned by comparison with a confidence threshold τ: ROIs with P(lateralized) ≥ τ were labelled as lateralized, those with P(lateralized) ≤ 1 − τ as terminal, and the remainder as indeterminate. The default value τ = 0.5 assigns a definite label to every ROI, since the two inequalities jointly cover the interval [0, 1] at that threshold; values τ > 0.5 introduce an indeterminate band centered at 0.5, trading coverage for classification confidence. For each ROI, the classifier output consisted of the bounding-box coordinates, the predicted class probabilities, and the assigned label.

2.6.3. Spatial Outputs and Probability Map Construction

For each tissue section, the inference module generated two complementary spatial representations of the predicted CX43 lateralization pattern (Figure 2E): a continuous probability map and a discrete classification overlay.

The continuous probability map was obtained by spatial interpolation of the ROI-level lateralization probabilities across the tissue section. For each classified ROI, the predicted probability P(lateralized) was added to an accumulator map P at the position of the ROI centroid (computed as the center of the bounding box); both P and a parallel weight map W, of the same spatial dimensions as the input image and initialized to zero, were used to track, respectively, the accumulated probability mass and the number of contributions received at each location. Both maps were then convolved with a two-dimensional Gaussian kernel G_σ of standard deviation σ, and the interpolated probability field $H(x,y)$ was computed as:

$$H(x,y) = \frac{(P * G_\sigma)(x,y)}{\max\{(W * G_\sigma)(x,y),\ \varepsilon\}}$$

where $*$ denotes two-dimensional convolution and $\varepsilon = 10^{-8}$ is a small constant that prevents division by zero in regions without contributing ROIs. This formulation corresponds to a Gaussian-weighted local average of the ROI-level probabilities and can be interpreted as a Nadaraya–Watson kernel estimator [30,31] applied to the sparse set of classified regions. The resulting field $H(x,y)$ represents, at each tissue location, a spatially smoothed estimate of the local lateralization probability derived from nearby classified regions: values near 1 correspond to tissue areas surrounded by ROIs predicted as lateralized with high confidence, while values near 0 correspond to areas dominated by terminal signal.

 doi:10.20944/preprints202605.0043.v1

The standard deviation σ of the Gaussian kernel determined the spatial scale at which the heatmap reflected variations in predicted lateralization. Rather than being tied to the absolute dimensions of the input image, σ was set based on the local density of detected ROIs, which is a property of the imaging conditions and tissue rather than of the image size. Specifically, σ was chosen so that the effective integration window of the kernel contained approximately k neighboring ROIs, following:

$$\sigma = \sqrt{\frac{k}{4\pi\rho}}$$

where ϱ is the ROI density, computed as the number of detected CX43-positive regions divided by the tissue area, and k is the expected number of ROIs contained within a disk of radius 2σ around any point of the tissue under the assumption of spatially uniform ROI distribution. The factor $4\pi\sigma^2$ corresponds to the effective integration area of the two-dimensional Gaussian kernel (i.e., the area of a disk of radius 2σ, which contains approximately 95% of the integrated kernel weight).

For the images analyzed in this study, acquired at 0.227 μm/pixel, the observed ROI density was of the order of 1.2×10^{-3} ROIs/μm² (equivalently, $\varrho = 6 \times 10^{-5}$ ROIs/px² at the native pixel resolution). Setting k = 12 for this density yielded σ ≈ 29 μm (126 px), which was adopted as the default configuration throughout this work. With this σ, the effective kernel area $4\pi\sigma^2$ spans approximately 10,500 μm², corresponding to a small local neighborhood that encompasses several adjacent cardiomyocytes given the typical dimensions reported for adult rat ventricular myocytes in histological sections (~100 μm in length and ~15–20 μm in width) [32] — a scale consistent with detecting tissue zones in which lateralization is shared among neighboring myocytes rather than being a per-cell phenomenon, and approximating the length scale over which direct intercellular coupling through gap junctions is expected to be functionally relevant. Smaller values of σ would produce heatmaps that largely replicate the discrete overlay, while much larger values would average out the spatial heterogeneity the map is intended to reveal.

When applying CLARISA to tissue sections acquired under different imaging conditions, σ was recomputed from the observed ROI density of the target image according to the equation above, keeping k fixed. In this scheme, k is the primary hyperparameter that characterizes the spatial scale of the heatmap in terms of the local ROI neighborhood, and σ is the resulting pixel-space standard deviation that realizes this scale on a given image. The numerical value of σ in pixels was the quantity actually supplied to the inference module, but this value is a deterministic function of k and the observed ROI density and carries no additional degree of freedom. Values of k in the range 10–15 are recommended: larger values produce smoother, more regionalized maps, while smaller values produce more localized maps that may retain individual ROI structure. Note that k affects only the continuous probability map and the complementary heatmap-based metrics defined in Supplementary Section S2.2; the primary area-based metrics are computed directly from the discrete per-ROI classification and are therefore independent of k.

The second output was a classification overlay, produced by filling the contour of each classified ROI with a color corresponding to its assigned hard label (terminal or lateralized). ROIs marked as indeterminate by the confidence rule described in Section 2.6.2 were rendered with a distinct color to make them identifiable at inspection. This overlay provided a direct, discrete visualization of the per-ROI classification that complements the continuous probability map. Both representations were exported as image files, and the continuous probability map was additionally saved as a color overlay blended with the original histological image.

2.6.4. Computation of Global Percent Lateralization

A global summary of CX43 lateralization at the tissue-section level was derived for each processed image. Let Ω_{all} denote the binary mask formed by the union of all detected ROI contours, let $\Omega_{conf} \subseteq \Omega_{all}$ denote the subset corresponding to ROIs assigned a confident label (i.e., excluding

 doi:10.20944/preprints202605.0043.v1

indeterminate ROIs), and let $\Omega_{lat} \subseteq \Omega_{conf}$ denote the subset corresponding to ROIs assigned a confident lateralized label. Two pairs of complementary metrics were computed.

The primary metrics quantified the fraction of the CX43-positive area assigned to the lateralized class were:

$$\%LatArea_{all} = 100 \times \frac{|\Omega_{lat|}}{|\Omega_{all}|}$$

$$\%LatArea_{conf} = 100 \times \frac{|\Omega_{lat|}}{|\Omega_{conf}|}$$

where $|\cdot|$ denotes pixel area. $\%LatArea_{all}$ expressed the lateralized fraction over the entire detected CX43-positive area and therefore penalizes indeterminate ROIs by including them in the denominator. By contrast, $\%LatArea_{conf}$ restricts the denominator to ROIs with a confident label, thus providing an estimate based only on regions for which the classifier committed to a decision. Both values were computed directly from the classification overlay described in Section 2.6.3 and therefore depended solely on the discrete per-ROI labels and the geometry of the detected contours, independently of the heatmap smoothing parameter *k*.

A complementary pair of heatmap-based metrics, $\%LatHeat_{all}$ and $\%LatHeat_{conf}$, was computed as the pixel-wise mean of $H(x,y)$ over the same masks. These metrics summarize the prediction in a confidence-weighted manner, allowing each pixel to contribute according to its propagated probability rather than according to a hard class label. Across all analyzed sections, the heatmap-based values remained within 5-6 percentage points of the corresponding area-based metrics. Their full definition and per-section values are reported in Supplementary Section S2.2.

2.7. Whole-Section Deployment: Evaluation Against Expert and Segmentation-Based Analysis

To illustrate the behavior of the inference pipeline described in Section 2.6 in a representative deployment scenario, image IM15 extracted from a different global image than those used during model development, and acquired at a slightly different spatial resolution (see Figure 5A), was analyzed by three independent sources: CLARISA, a domain expert working manually, and a previously published segmentation-based methodology (MARTA) [18]. As noted in Section 2.1, this image was acquired at a spatial resolution slightly different from that of the development set, so that the evaluation also probes the robustness of the pipeline to acquisition conditions not seen during training.

Because IM15 was acquired at a slightly lower spatial resolution than the development set (0.3 μm/pixel vs. 0.227 μm/pixel), the resolution-dependent parameters of the inference pipeline were adjusted accordingly: the local and context crop sizes were rescaled from 256 × 256 and 512 × 512 to 192 × 192 and 384 × 384 pixels, preserving the physical field of view captured by each crop (Section 2.6.2). The neighborhood hyperparameter *k* was kept at its default value of 12, yielding a kernel standard deviation of σ = 94 px for this image (compared with σ = 126 px for the development-resolution images on the test slides), as derived from the formula given in Section 2.6.3 and the observed ROI density of each image.

In addition, the three parameters governing ROI detection were adjusted to accommodate the imaging characteristics of this image, which exhibited weaker and more spatially fragmented CX43 signal than the development-set images. First, *thresh_value* was lowered from 180 to 150 so that faintly stained CX43-positive regions, which would have fallen below the default threshold, were preserved in the binary mask. Second, *kernel_open* was reduced from 9 to 5 because, under weak staining, small isolated detections are more likely to correspond to genuine low-intensity CX43 signal than to background noise, and the default kernel size would have eliminated them irreversibly. Third, *kernel_dilate* was increased from 5 to 21 so that the fragmented CX43-positive regions — which appeared as multiple small detections separated by dark gaps rather than as compact structures — could grow and reconnect into ROIs representing the underlying biological regions as coherent units (see Table 4). All three adjustments therefore respond to a single underlying feature of the evaluation

 doi:10.20944/preprints202605.0043.v1

image: a weaker and more fragmented CX43 signal than that present in the development set. The decision threshold τ was kept at its default value of 0.5.

The same image was independently analyzed by a domain expert — the same individual who generated the training dataset described in Section 2.2 — using a custom interactive tool developed in-house for this study, which presented every CX43-positive region detected by the automatic ROI pipeline (Section 2.6.1) and allowed the expert to assign a class label to each of them; additional details on the interface are provided in Supplementary Section S1.3. For each region, the expert assigned one of three labels: terminal, lateralized, or uncertain, the last being used when the region could not be confidently assigned to either class. Expert-derived spatial outputs and global metrics were then computed by feeding these manual classifications into the same functions of the inference pipeline (Sections 2.6.3 and 2.6.4), with each label converted to a binary probability (P(lateralized) = 1, 0, or 0.5, respectively) as input to the spatial interpolation. This route is representative of the current manual alternative available to a researcher who wishes to quantify lateralization on a given section without relying on automated classification.

IM15 was additionally processed with the previously published segmentation-based methodology (MARTA) [18], applied in its original form without re-optimization for the imaging conditions used in this study. This route is representative of the previous automated alternative against which CLARISA is most directly comparable.

Because the three sources operate on partially different populations of CX43 signal, only the area-based metrics defined in Section 2.6.4 ($\%LatArea_{all}$ and $\%LatArea_{conf}$) were used for numerical comparison. Heatmap-based metrics were not considered in this comparison, as they are not meaningful when derived from binary expert labels. Inference with CLARISA was performed at the default decision threshold ($\tau = 0.5$) without applying a confidence margin; under this configuration no ROIs are assigned to the indeterminate category, and $\%LatArea_{conf}$ therefore coincides by construction with $\%LatArea_{all}$ and is not reported separately for CLARISA.

Table 4. Inference parameters used for the deployment evaluation. Resolution-dependent parameters (crop sizes) were rescaled according to the formula given in Section 2.6.2. The neighborhood hyperparameter k was kept at its default value; the corresponding kernel standard deviation σ, reported here for reference, is a derived quantity that was recomputed from the observed ROI density of each image (Section 2.6.3). The three detection parameters were jointly adjusted in response to the weaker and more spatially fragmented CX43 signal observed on the evaluation section.

Parameter	Development set	IM15	Reason for adjustment
Spatial resolution	0.227 µm/pixel	0.3 µm/pixel	Acquisition condition
crop_local	256 px	192 px	Rescaled to preserve physical FOV (Section 3.6.2)
crop_context	512 px	384 px	Rescaled to preserve physical FOV (Section 3.6.2)
σ (derived, in px)	126 px	94 px	Recomputed from observed ROI density (Section 3.6.3)
thresh_value	180	150	Weaker CX43 staining on the evaluation section
kernel_open	9	5	Reduced to avoid eliminating genuine low-intensity signal
kernel_dilate	5	21	Increased to reconnect spatially fragmented CX43-positive regions
τ (decision threshold)	0.5	0.5	Default retained

 doi:10.20944/preprints202605.0043.v1

3. Results

3.1. Architecture and Input Representation Optimization

Optimization was performed in two sequential stages: an architecture search followed by hyperparameter fine-tuning.

3.1.1. Architecture Search Identified the Importance of Tissue Context

The first optimization stage showed that local-only models consistently underperformed architectures incorporating broader tissue context, indicating that the appearance of the CX43-positive focus alone was insufficient for robust discrimination between terminal and lateralized patterns. By contrast, both contextual-only and dual-scale formulations achieved better validation performance. The lowest validation loss in this stage was obtained by a contextual-only MLP model (validation loss = 0.0451), whereas the highest validation AUC was achieved by a dual-scale MLP model (validation AUC = 0.8389), which also retained similarly low validation loss. In addition, MLP-based heads consistently outperformed logistic heads among the top-ranked configurations, supporting the use of a non-linear decision layer. Taken together, these results indicate that successful classification benefits from both local morphology and tissue context. Accordingly, the dual-scale MLP architecture was selected as the base model for subsequent fine-tuning, as it combined competitive validation loss with the highest discriminative performance. Full trial-wise results are provided in Supplementary File S1 (architecture).

3.1.2. Fine-Tuning Optimization Further Improved Validation Performance

After fixing the base architecture, a second Optuna study was performed to optimize training-related hyperparameters. This stage further improved validation performance, reducing the best validation loss by approximately 10.4% and increasing validation AUC from 0.832 to 0.862. The top-performing configurations consistently favored stronger regularization and a longer final fine-tuning stage, suggesting that performance benefited from extended joint adaptation of the full network after the earlier staged specialization steps. These configurations also supported the use of discriminative learning rates, with larger effective updates in the classification head and more conservative updates in the pretrained backbone. In addition, the best checkpoints were typically reached during stage 3, when the full network was unfrozen. Together, these results indicate that, beyond architectural design, classifier performance depended on progressively adapting the pretrained backbone to the morphological characteristics of CX43-positive regions. Full trial-wise results are provided in Supplementary File S1 (fine-tuning).

3.2. Final Model Performance on Held-Out Slides

The final model was evaluated on the held-out test slides IM1313 and IM1314, providing an assessment of generalization to previously unseen tissue sections. Overall, the model performed well on the terminal class, with somewhat lower performance on the lateralized class, reaching a ROC-AUC of 0.905 and a PR-AUC of 0.810 on the combined test set (see Figure 3A).

When performance was examined at the slide level, a more heterogeneous pattern emerged (see Figure 3B. In IM1313, the model showed a highly conservative behaviour for the lateralized class: no false positives were observed and longitudinal ROIs were recovered with perfect recall, but only one of five lateralized ROIs were correctly identified (Figure 3B – left). By contrast, IM1314 displayed a more balanced behaviour, with 15 of 19 lateralized ROIs correctly classified, corresponding to a lateralized-class precision of 0.714 and recall of 0.789, although 6 false positives were also observed (Figure 3B – right). Thus, model performance was not homogeneous across the held-out test set, but depended strongly on the slide being analyzed.

At the ROI level, the distribution of predicted probabilites further showed that correct predictions tended to occupy more extreme probability ranges, whereas misclassifications were

 doi:10.20944/preprints202605.0043.v1

enriched closer to the decision threshold (see Figure 3C – left). A similar slide-dependent pattern was also observed in the training and validation sets (Supplementary Figures S3–S5), suggesting that ROI-level classification difficulty is influenced not only by slide-specific morphological context, but also by the presence of visually ambiguous regions that may be challenging even at the annotation level. These cases are examined in more detail in the following Discussion section.

Given the limited number of ROIs available in the test set, these estimates should nevertheless be interpreted with appropriate caution. Additional performance details for the final retrained model on the training and validation sets are provided in Supplementary Section S2.1.

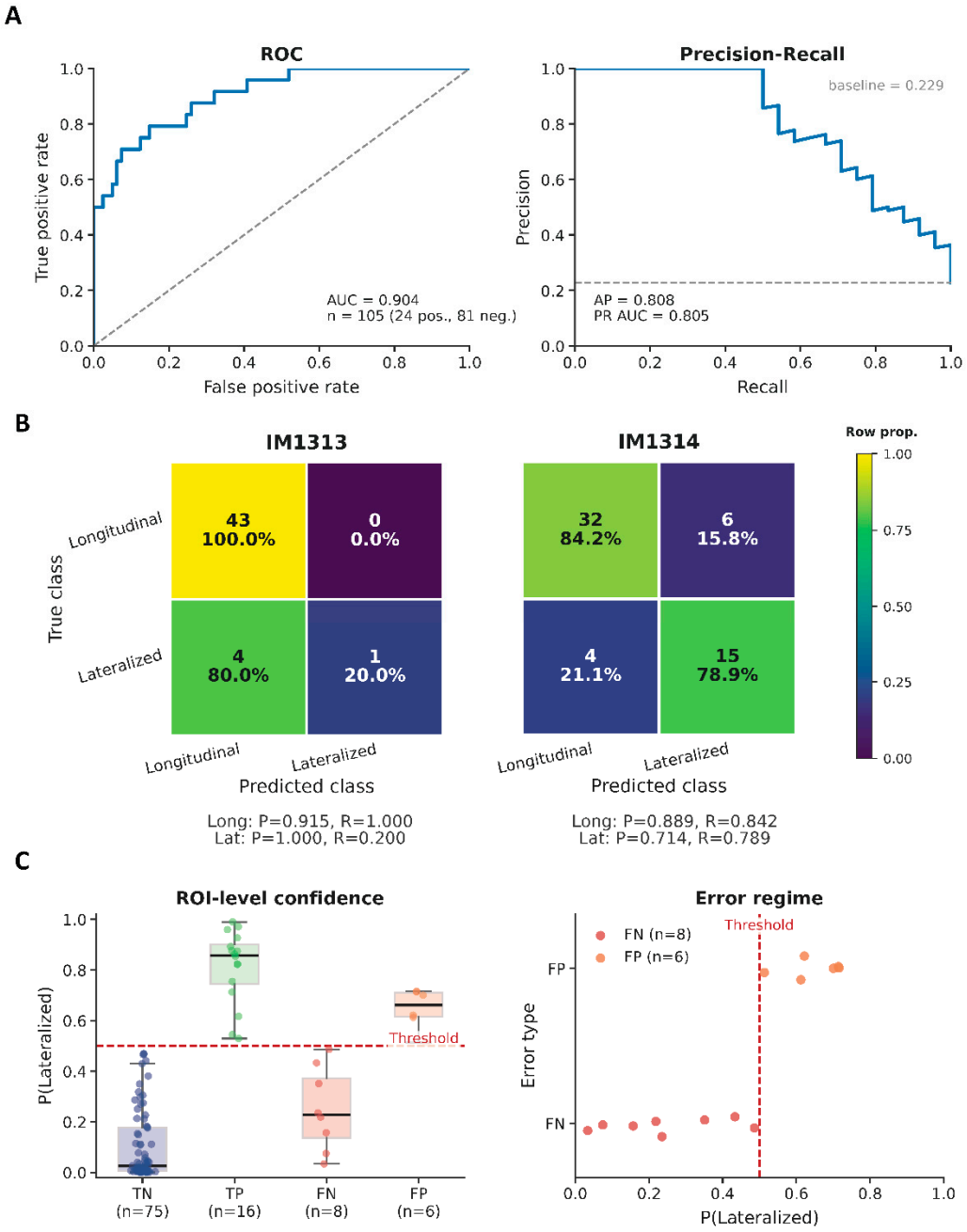

Figure 3. Quantitative performance of the final MARTA classifier on the held-out test sections. (A) Global discrimination performance on the combined test set, shown as ROC and precision–recall curves. (B) Slide-level confusion matrices for the held-out test slides IM1313 and IM1314. Rows correspond to the true class and columns to the predicted class. Cell colors indicate row-normalized proportions, allowing direct comparison of classification behaviour across slides despite differences in class distribution. Each cell reports both the absolute

 doi:10.20944/preprints202605.0043.v1

number of ROIs and the corresponding within-row percentage. Precision and recall for the terminal (longitudinal) and lateralized classes are reported below each panel. (C) ROI-level confidence profiles in the test set. Left, distribution of predicted probabilities for the lateralized class grouped by prediction outcome (TN, TP, FN, FP). Right, predicted probabilities for false negatives and false positives relative to the default decision threshold of 0.5. Together, these panels show that correct predictions tend to occupy more extreme probability ranges, whereas misclassifications are enriched near the decision boundary.

3.3. Whole-Section Inference on Held-Out Slides

To examine whole-section behavior of the inference workflow described in Section 2.6, the trained classifier was applied to two complete tissue sections from the held-out test set (IM1313 and IM1314). In contrast to the ROI-level evaluation reported above, which was restricted to expert-annotated regions, this analysis included all CX43-positive regions automatically detected across each section. The number of regions processed was therefore substantially larger than the number of annotated ROIs contributed by these sections to the held-out test set (Section 2.5.1), corresponding to the full CX43-positive signal identifiable by the detection pipeline. A total of 3,715 regions were processed in IM1313 and 1,233 in IM1314.

For each section, the inference module generated the complete set of outputs defined in Sections 2.6.3 and 2.6.4, including a continuous probability map of predicted lateralization, a classification overlay assigning each detected ROI to one of the two classes, and the global lateralization metrics. The Figure 4 shows the original fluorescence image (panels A and D), the continuous lateralization heatmap (panels B, E), and the distribution of ROI-level predicted probabilities partitioned by assigned class (panels C, F) for each tissue section. In both sections, the heatmaps indicated that predicted lateralization was not uniformly distributed across the tissue. Instead, localized regions with higher predicted lateralization probability were observed, particularly in IM1314, whereas IM1313 showed a more homogeneous overall pattern with fewer and more scattered high-probability foci.

The class-partitioned probability distributions (Figure 4C-F) showed separation between ROIs assigned to the terminal and lateralized classes in both sections. In IM1313, ROIs assigned to the terminal class concentrated near zero probability values (median P(lateralized) = 0.001, Q1–Q3: 0.000–0.019), whereas ROIs assigned to the lateralized class were concentrated at higher probabilities (median P(lateralized) = 0.759, Q1–Q3: 0.618–0.883). In IM1314, a similar pattern was observed, with somewhat broader distributions (terminal: median 0.017, Q1–Q3: 0.001–0.132; lateralized: median 0.875, Q1–Q3: 0.698–0.968). Only a small fraction of ROIs fell near the decision threshold: in IM1313, 3.15% of ROIs had probabilities between 0.3 and 0.5, and in IM1314, 8.27%. Under the default binary decision rule ($\tau = 0.5$), every ROI was assigned to one of the two classes; accordingly, the *all* and *conf* variants of the global metrics were identical for both sections.

The global whole-section metrics differed markedly between the two tissue sections. In IM1313, 3,471 ROIs were classified as terminal and 244 as lateralized, yielding $\%LatArea_{all} = 6.63$. In IM1314, 948 ROIs were classified as terminal and 285 as lateralized, yielding $\%LatArea_{all} = 19.06$. Thus, the proportion of predicted lateralized signal was substantially higher in IM1314 than in IM1313 (see Table 5).

Table 5. Whole-section inference results on held-out test sections.

Section	N detected	Terminal	Lateralized	% $LatArea_{all}$
IM1313	3,715	3,471	244	6.63
IM1314	1,233	948	285	19.06

 doi:10.20944/preprints202605.0043.v1

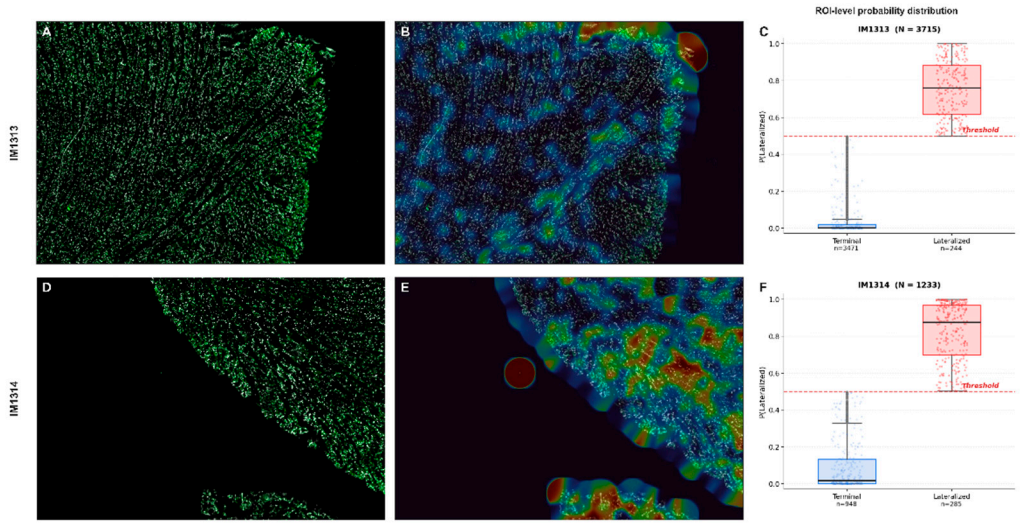

Figure 4. Whole-section inference outputs of the trained classifier on the two held-out test sections IM1313 (top row) and IM1314 (bottom row). (A, D) Original fluorescence images used as input for inference. (B, E) Continuous heatmaps of predicted lateralization generated from ROI-level probabilities by Gaussian kernel interpolation ($k = 12$, yielding $\sigma \approx 29$ μm). Cooler colors indicate lower predicted lateralization, whereas warmer colors indicate higher predicted lateralization. (C, F) Boxplots of predicted lateralization probability (P(lateralized)) at the ROI level, partitioned by assigned class (terminal, P(lateralized) < 0.5; lateralized, P(lateralized) ≥ 0.5). Boxes show median and interquartile range; whiskers extend to 1.5× IQR; individual ROIs are shown as jittered points (subsampled for visualization when n > 300). The default decision threshold (0.5) is indicated by the dashed red line. The number of ROIs per class is reported below each box.

3.4. Whole-Section Inference: Comparison with Expert Annotations and with a Segmentation-Based Methodology

CLARISA detected 359 CX43-positive ROIs. All of them received a confidence-passing label, so no region was assigned to the indeterminate class: because the decision threshold was fixed at $\tau = 0.5$, every ROI falls on one side or the other by construction. The expert independently labelled 185 of these ROIs as lateralized, 160 as terminal and 14 as uncertain.

The continuous lateralization heatmaps produced by the expert annotations and by CLARISA showed close overall agreement in spatial structure (see Figure 5B–C).

Both representations identified the same principal warm band, running vertically along the central portion of the section, as well as the same cold regions along the upper-left, lower-left and upper-right borders, so that the dominant spatial pattern of lateralization — a predominantly lateralized central axis flanked by predominantly terminal lateral zones — was preserved across the two sources. Two consistent differences were nevertheless observed. First, the expert-derived heatmap reached more extreme values locally, most visibly in a saturated warm focus in the lower-central region where CLARISA produced a warm but less intense response; conversely, CLARISA produced a broader warm zone along the mid-left margin than the expert. Second, the expert-derived heatmap was intrinsically coarser and exhibited sharper transitions between neighbouring zones, whereas CLARISA produced smoother gradients with more diffuse boundaries. This textural difference is a direct consequence of the reconstruction process: the expert heatmap is interpolated from binary labels (with uncertain regions fixed at 0.5), while the classifier outputs continuous probabilities, and the smoother appearance of the latter should therefore not be read as indicative of better or worse agreement with the underlying biology.

At the ROI level, CLARISA assigned systematically different probability distributions across the two expert-annotated categories (Figure 5C). For ROIs annotated as terminal, predicted probabilities were concentrated toward lower values, with a median of 0.24 and an interquartile range of 0.09-0.62.

 doi:10.20944/preprints202605.0043.v1

For ROIs annotated as lateralized, predicted probabilities were shifted toward higher values, with a median of 0.64 and an interquartile range of 0.40-0.88. Thus, the central tendency of the two distributions fell on opposite sides of the 0.5 decision threshold, although partial overlap remained around the threshold. In particular, some ROIs annotated as terminal showed predicted probabilities above 0.5, whereas some annotated as lateralized showed predicted probabilities below 0.5. Fourteen ROIs in the whole section had been labeled as uncertain by the expert and were excluded from this boxplot analysis (see Figure 6).

At the tissue-section level, CLARISA yielded $\%LatArea_{all}$ = 44.26%. The corresponding values derived from expert annotations were 42.96% for $\%\ LatArea_{all}$ and 45.20% for $\%\ LatArea_{conf}$, the difference between the two expert values reflecting the ROIs that the expert marked as uncertain. The absolute difference between the CLARISA and the expert was 1.30 percentage points when compared against $\%\ LatArea_{all}$ and 0.95 percentage points when compared against $\%\ LatArea_{conf}$. Both summaries were computed over the same total detected CX43-positive area (100%), so the comparison is not affected by differences in signal coverage, but only by how the detected ROIs were assigned to terminal, lateralized, or uncertain categories.

When the same section was analyzed with MARTA, quantification was restricted to the fraction of CX43-positive signal contained within successfully segmented cardiomyocytes. In this image, MARTA segmented 82 cardiomyocytes, which did not correspond to the full cardiomyocyte population visible in the section (Figure 5D). As a result, part of the CX43-positive signal lay outside any successfully segmented cell and was not included in MARTA's analysis. MARTA output was also represented differently from that of CLARISA: rather than assigning a class directly to each detected CX43-positive ROI, MARTA displayed each segmented cardiomyocyte as an oriented bounding box subdivided into four longitudinal compartments along the principal cellular axis, as defined in the original methodology [18]. The two outermost compartments corresponded to terminal regions and the two central compartments to lateral regions; CX43 signal was then quantified according to the compartment in which it fell. Within its analysis population, MARTA yielded a global percent lateralization of 22.18%.

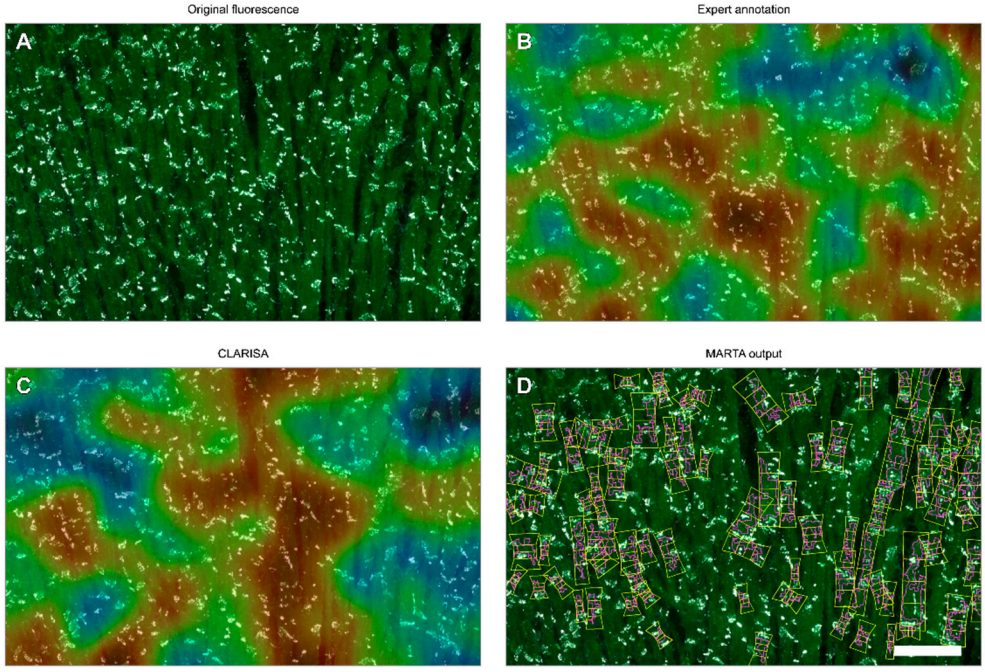

Figure 5. Whole-section comparison between CLARISA, expert annotations, and the MARTA segmentation-based methodology. (A) Original fluorescence image of the analyzed tissue section, used as reference for all visual outputs. (B) Continuous lateralization heatmap derived from the expert annotations, computed by feeding

 doi:10.20944/preprints202605.0043.v1

the binary expert labels into the same interpolation procedure used for panel C. Regions labelled as uncertain by the expert were fixed at a probability of 0.5 before interpolation. (C) Continuous lateralization probability heatmap produced by CLARISA, computed as described in Section 2.6.3. (D) Output of MARTA applied to the same tissue section. Each successfully segmented cardiomyocyte is shown with its oriented bounding box subdivided into four longitudinal compartments along the principal cellular axis, as defined in the original methodology: the two outermost compartments represent the terminal regions of the cell, while the two central compartments represent the lateral regions. CX43 signal falling within each compartment is assigned to the terminal or lateral class accordingly. Tissue regions where no cardiomyocyte was successfully segmented remain unshaded, illustrating the coverage gap discussed in the text. Panels B and C share the same colormap and value range [0, 1].

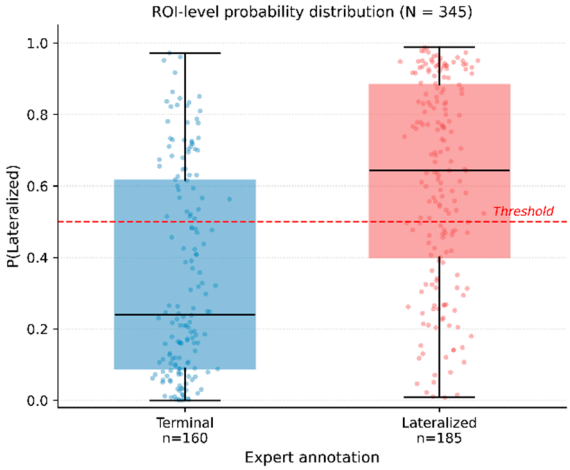

Figure 6. ROI-level distribution of CLARISA-predicted P(lateralized) grouped by expert annotation. Boxplots show the CLARISA-predicted probability of lateralization for the 345 CX43-positive ROIs with a confident expert label, separated according to whether the expert annotated them as terminal (blue, n = 160) or lateralized (red, n = 185). Boxes indicate the interquartile range and median; whiskers extend to the most extreme non-outlier values. Individual ROIs are overlaid with horizontal jitter. The red dashed line marks the P(lateralized) = 0.5 decision threshold used by CLARISA.

4. Discussion

4.1. Limitations Related to Training Data Diversity and External

A key limitation of the present study lies in the restricted diversity of the development dataset. Although the classifier was trained using transfer learning and a relatively large number of annotated CX43-positive ROIs, these annotations originated from only two global images and a limited number of tissue sections. As a result, the training data may not fully capture the broader variability that can affect CX43 appearance in practice, including differences in section orientation with respect to the myocardial fibers, staining intensity, signal fragmentation, image contrast, blur, or acquisition conditions.Optimization was performed in two sequential stages: an architecture search followed by hyperparameter fine-tuning.

Nevertheless, CLARISA is intended not only as a fixed classifier trained on the present dataset, but also as a practical framework that can be used either directly with the pretrained weights provided here or adapted to new experimental settings through additional annotation and retraining. For this purpose, the public repository includes both the full training pipeline and an expert annotation tool based on the automatic ROI detection procedure described in Supplementary Section S1.3, designed to support faster and more efficient expert labeling when expanding the current dataset or generating a new one under different experimental conditions. In the present work, this

 doi:10.20944/preprints202605.0043.v1

tool was used to annotate the IM15 section, which served as the basis for the comparison between expert visual inspection and MARTA output (see Sections 2.7 and 3.4).

4.2. Spatial Mixed Error Patterns and Local Ambiguity

With respect to model behavior, spatial inspection of the predictions on test tissue slides IM1313 and IM1314 suggests that errors are not organized into large, homogeneous regions of failure. Instead, both slides show locally mixed neighborhoods in which correctly and incorrectly classified ROIs coexist (see Figure 7). Thus, the main message of the spatial analysis is not the presence of fully misclassified image regions, but rather the existence of locally ambiguous microenvironments.

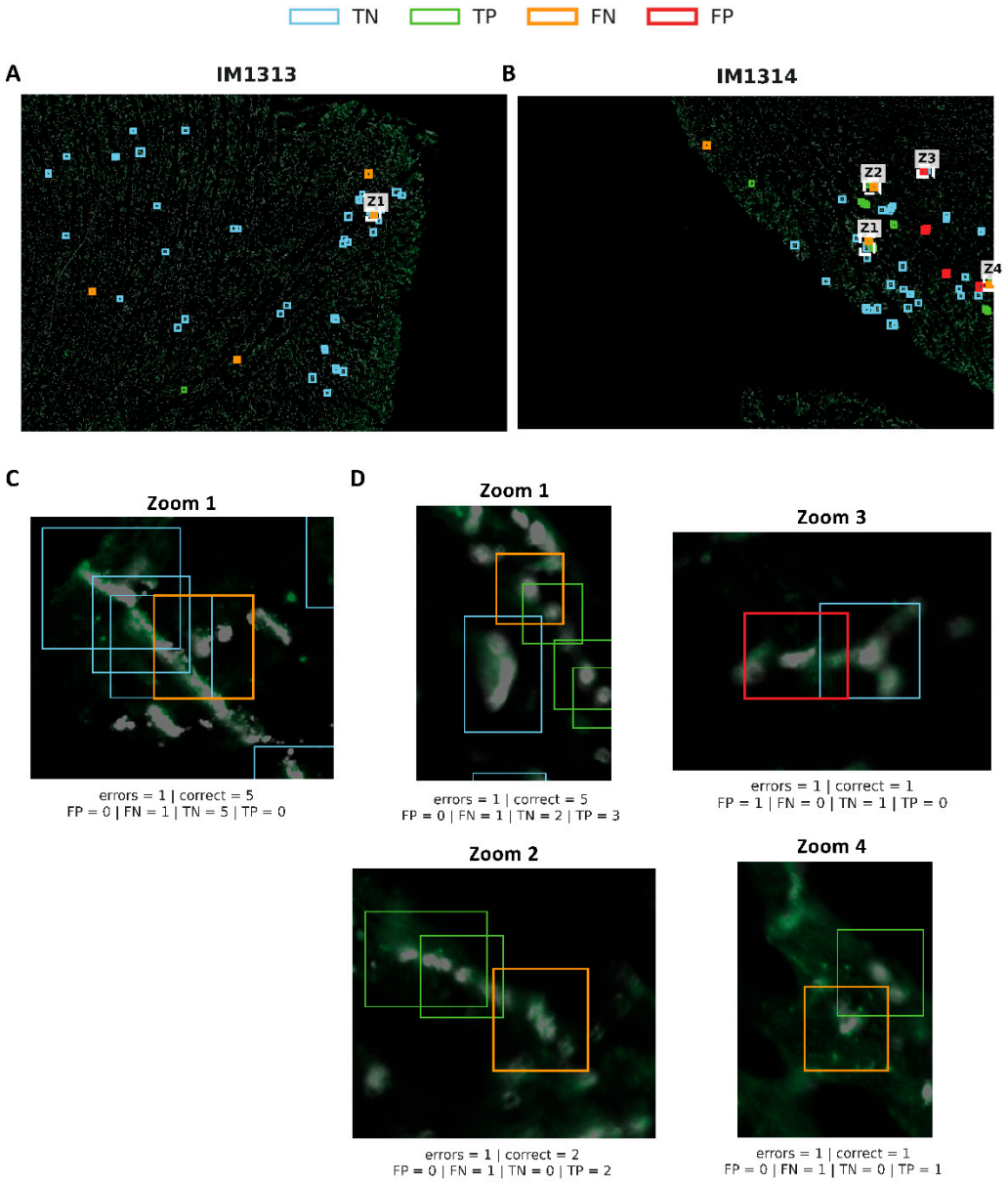

Figure 7. Spatial distribution of ROI-level prediction outcomes in held-out slides IM1313 and IM1314. ROI-level predictions were mapped back to their original whole-slide coordinates and classified as true negatives (TN, blue), true positives (TP, green), false negatives (FN, orange), and false positives (FP, red). Global views of IM1313 and IM1314 are shown in A and B, respectively. Zoomed regions (C and D) were selected to highlight locally informative neighborhoods containing nearby or partially overlapping ROIs with concordant or discordant outcomes. Rather than forming large homogeneous blocks of failure, misclassifications were typically observed in locally mixed regions in which correct and incorrect predictions coexisted. Some partially

 doi:10.20944/preprints202605.0043.v1

overlapping ROIs nevertheless remained correctly classified, suggesting tolerance to a certain degree of sample overlap within the larger input crops. ROI identifiers and bounding-box coordinates corresponding to the zoomed examples are listed in Supplementary Table S5.

To examine this pattern more closely, we selected regions displaying local coexistence of correct and incorrect predictions, as well as high proximity or partial overlap between ROIs (see Figure 7C-D). The zoomed views show that ROIs located very close to one another, and in some cases partially overlapping, can nevertheless yield different pre-dictions, pointing to sensitivity to fine spatial differences in the visual content being analyzed. At the same time, it is important to note that the bounding box displayed on the slide in Figure 7 does not exactly match the full 256/512-pixel crop used as input to the model. Therefore, spatial proximity or partial overlap between bounding boxes does not necessarily mean that the model received identical visual information. Even so, the fact that some partially overlapping ROIs remain correctly classified indicates that shared content between neighboring crops is not, by itself, sufficient to produce sys-tematic failure. Taken together, these observations suggest that the model's limitations are more closely related to fine-scale local ambiguity than to an inability to recognize larger image regions as a whole.

This interpretation is further supported by the representative examples shown in Supplementary Section S3. Overall, the 256- and 512-pixel crops indicate that many discordant predictions arise in morphologically challenging or borderline regions, whereas clearly correct cases tend to be associated with extreme lateral probabilities, consistent with high model confidence when the underlying pattern is well defined. Several of the observed discrepancies are compatible with transitional or borderline regions between longitudinal and lateralized patterns, as suggested by examples such as ROI 3012 in IM1313 and ROIs 3028 and 3034 in IM1314 (see Supplementary Figure S7-S8). Other discordant cases are associated with blur, low-contrast or atypical signal, as illustrated by examples including ROIs 3011, 3013, 3055, and 3063 Supplementary Figure S7-S8). In selected cases, such as ROI 3069 (see Supplementary Figure S8-B, difficulties in centering the analysis window may also contribute. This reflects a methodological limitation of the crop-generation procedure: when the bounding box lies close to the image border, the corresponding 256- or 512-pixel crop cannot always be centered optimally around the signal of interest. In this context, part of the observed error may reflect not only limitations of the model itself, but also an intrinsic interpretative difficulty of the task.

4.3. Whole-Section Inference: Deployment, Comparison, and Methodological Considerations

4.3.1. Deployment on Held-Out Whole Sections

The two held-out sections exhibited distinct whole-section spatial patterns. IM1314 showed more prominent clustered high-probability regions and a higher global lateralization estimate than IM1313, whereas IM1313 displayed a more homogeneous pattern with fewer and more scattered lateralized foci. These observations indicate that CLARISA can generate spatially structured outputs at the tissue-section scale, rather than merely producing independent ROI-level predictions.

The ROI-level probability distributions also differed between the two held-out sections. In IM1313, ROIs assigned to the terminal class were tightly concentrated near zero, whereas ROIs assigned to the lateralized class showed a broader spread of predicted probabilities above the decision threshold. By contrast, in IM1314 the lateralized class appeared more compact at high probability values, while the terminal class was more broadly distributed. This section-specific difference suggests that the confidence structure of CLARISA's predictions was not identical across the two images. In particular, the broader dispersion of lateralized predictions in IM1313 is consistent with the ROI-level ambiguities previously observed for samples derived from this section in Section 4.2. Although this correspondence should not be overinterpreted from whole-section distributions alone, it suggests that IM1313 may contain a larger proportion of regions lying closer to the decision boundary and may therefore merit closer examination.

 doi:10.20944/preprints202605.0043.v1

This analysis should be interpreted primarily as a demonstration of whole-section deployment rather than as a formal validation of spatial accuracy. Because no exhaustive expert annotation was available at the section level for these held-out images, its main value lies in showing that CLARISA remains operational, interpretable, and quantitatively coherent when applied to complete tissue sections.

4.3.2. Comparison with Expert Annotation and MARTA on a Novel Whole Section

Whereas Section 3.3 focused on whole-section deployment in held-out test sections, Section 3.4 examined CLARISA on a genuinely novel image by comparing its outputs with expert annotation and with the previously proposed MARTA methodology. At the tissue-pattern level, the expert-derived and CLARISA-derived heatmaps showed similar large-scale spatial organization, including a central warm band and cooler peripheral regions. At the same time, the agreement was not exact. The two heatmaps differed locally in the extent and intensity of some warm and cool regions, and the ROI-level probability distributions showed overlap around the 0.5 threshold, indicating that a subset of detected regions received different assignments from CLARISA and the expert.

At the ROI level, CLARISA generally assigned lower predicted lateralization probabilities to ROIs annotated by the expert as terminal and higher probabilities to those annotated as lateralized. This separation is consistent with agreement between model output and expert labeling at the level of individual detected regions. At the same time, the overlap observed around the decision threshold indicates that not all ROIs were cleanly separable into the two categories. Rather than being interpreted solely as disagreement, this pattern suggests that a subset of regions lies near the decision boundary and may be intrinsically ambiguous, which is consistent with the fact that the expert identified a subset of ROIs in the same tissue section as uncertain.

Comparison with MARTA highlights a methodological difference relevant to whole-section applicability. MARTA depends on successful cardiomyocyte segmentation and therefore quantifies only the subset of CX43-positive signal that falls within segmented cells. In the present section, this constraint limited its effective analysis domain, as illustrated by the restricted coverage visible in Figure 5D. By contrast, CLARISA operates directly on detected CX43-positive regions and is therefore not constrained by incomplete cell segmentation. This difference is reflected in the section-level estimates: CLARISA was markedly closer to the expert-derived $\%LatArea_{all}$ value than MARTA, with absolute differences of 1.30 and 20.78 percentage points, respectively. Accordingly, the discrepancy should not be interpreted only as a numerical difference, but also in light of the effective analysis domain available to each method. In this section, MARTA's dependence on successful cardiomyocyte segmentation substantially reduced the fraction of CX43-positive signal that could be quantified.

5. Conclusions

In this study, we developed CLARISA, a framework for the automated analysis of CX43 lateralization in fluorescence images of ventricular myocardium without requiring explicit cardiomyocyte segmentation. To support this approach, we generated an expert-annotated dataset of CX43-positive ROIs classified as terminal or lateralized.

Based on this dataset, we trained a transfer-learning classifier to predict the lateralization status of individual CX43-positive regions and used it as the core of an inference pipeline for whole tissue sections, enabling estimation of the proportion and spatial distribution of lateralized CX43 signal across complete images.

Beyond the classifier itself, CLARISA is provided as a practical and extensible framework: the trained model, together with the expert annotation and training tools, is publicly available, allowing users either to apply the pretrained workflow to similar data or to generate their own datasets and adapt the classifier to different tissue preparations, species, staining protocols, or imaging conditions.

Overall, CLARISA is intended as a practical and scalable tool for studying CX43 lateralization across different cases and experimental settings, particularly in analyses where manual quantification would be labor-intensive and difficult to scale. By automating this process, it enables a more

 doi:10.20944/preprints202605.0043.v1

systematic and efficient assessment of lateralization patterns and may facilitate their study across broader cohorts and pathological contexts.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org.

References

1. Spach, M.S.; Heidlage, J.F. The stochastic nature of cardiac propagation at a microscopic level. Electrical description of myocardial architecture and its application to conduction. Circ. Res. 1995, 76, 366–380. doi:10.1161/01.RES.76.3.366.
2. Kléber, A.G.; Rudy, Y. Basic mechanisms of cardiac impulse propagation and associated arrhythmias. Physiol. Rev. 2004, 84, 431–488. doi:10.1152/physrev.00025.2003.
3. Dhein, S.; Seidel, T.; Salameh, A.; Jozwiak, J.; Hagen, A.; Kostelka, M.; Hindricks, G.; Mohr, F.W. Remodeling of cardiac passive electrical properties and susceptibility to ventricular and atrial arrhythmias. Front. Physiol. 2014, 5, 424. doi:10.3389/fphys.2014.00424.
4. Beňová, T.E.; Baˇcová, B.S.; Viczenczová, C.; Diez, E.; Barancik, M.; Tribulová, N. Protection of cardiac cell-to-cell coupling attenuate myocardial remodeling and proarrhythmia induced by hypertension. Physiol. Res. 2016, 65 (Suppl. 1), S29–S42. doi:10.33549/physiolres.933391.
5. Prado, N.J.; Egan Beňová, T.; Diez, E.R.; Knezl, V.; Lipták, B.; Ponce Zumino, A.Z.; Llamedo-Soria, M.; Szeiffová Baˇcová, B.; Miatello, R.M.; Tribulová, N. Melatonin receptor activation protects against low potassium-induced ventricular fibrillation by preserving action potentials and connexin-43 topology in isolated rat hearts. J. Pineal Res. 2019, 67, e12605. doi:10.1111/jpi.12605.
6. Severs, N.J.; Bruce, A.F.; Dupont, E.; Rothery, S. Remodelling of gap junctions and connexin expression in diseased myocardium. Cardiovasc. Res. 2008, 80, 9–19. doi:10.1093/cvr/cvn133.
7. Seidel, T.; Salameh, A.; Dhein, S. A simulation study of cellular hypertrophy and connexin lateralization in cardiac tissue. Biophys. J. 2010, 99, 2821–2830. doi:10.1016/j.bpj.2010.09.010.
8. Winters, J.; Edler von Braunmuhl, M.; Zeemering, S.; Gilbers, M.; Ten Brink, T.; Scaf, B.; Guasch, E.; Mont, L.; Batlle, M.; Sinner, M.; et al. JavaCyte, a novel open-source tool for automated quantification of key hallmarks of cardiac structural remodeling. Sci. Rep. 2020, 10, 20074. doi:10.1038/s41598-020-76932-3.
9. Greiner, J.; Frangiamore, F.; Sonak, F.; Madl, J.; Seidel, T.; Kohl, P.; Rog-Zielinska, E.A. A deep learning-enabled toolkit for the 3D segmentation of ventricular cardiomyocytes. J. Physiol. 2025. doi:10.1113/JP288557.
10. da Silva, F.S.; de Souza, N.C.S.A.; de Moraes, M.V.; Abreu, B.J.; de Oliveira, M.F. CmyoSize: An ImageJ macro for automated analysis of cardiomyocyte size in images of routine histology staining. Ann. Anat. 2022, 241, 151892. doi:10.1016/j.aanat.2022.151892.
11. Peyster, E.G.; Madabhushi, A.; Margulies, K.B. Advanced morphologic analysis for diagnosing allograft rejection: the case of cardiac transplant rejection. Transplantation 2018, 102, 1230–1239. doi:10.1097/TP.0000000000002189.
12. Karpurapu, A.; Williams, H.A.; DeBenedittis, P.; Baker, C.E.; Ren, S.; Thomas, M.C.; Beard, A.J.; Devlin, G.W.; Harrington, J.; Parker, L.E.; et al. Deep learning resolves myovascular dynamics in the failing human heart. JACC Basic Transl. Sci. 2024, 9, 674–686. doi:10.1016/j.jacbts.2024.02.007.
13. Lipkova, J.; Chen, T.Y.; Lu, M.Y.; Chen, R.J.; Shady, M.; Williams, M.; Wang, J.; Noor, Z.; Mitchell, R.N.; Turan, M.; et al. Deep learning-enabled assessment of cardiac allograft rejection from endomyocardial biopsies. Nat. Med. 2022, 28, 575–582. doi:10.1038/s41591-022-01709-2.
14. Lackey, D.P.; Carruth, E.D.; Lasher, R.A.; Boenisch, J.; Sachse, F.B.; Hitchcock, R.W. Three-dimensional modeling and quantitative analysis of gap junction distributions in cardiac tissue. Ann. Biomed. Eng. 2011, 39, 2683–2694. doi:10.1007/s10439-011-0369-3.
15. Yan, J.; Thomson, J.K.; Wu, X.; Zhao, W.; Pollard, A.E.; Ai, X. Novel methods of automated quantification of gap junction distribution and interstitial collagen quantity from animal and human atrial tissue sections. PLoS ONE 2014, 9, e104357. doi:10.1371/journal.pone.0104357.

 doi:10.20944/preprints202605.0043.v1

16. Hofgaard, J.P.; Mollerup, S.; Holstein-Rathlou, N.H.; Nielsen, M.S. Quantification of gap junctional intercellular communication based on digital image analysis. Am. J. Physiol. Regul. Integr. Comp. Physiol. 2009, 297, R857–R867. doi:10.1152/ajpregu.00089.2009.
17. Boyle, P.M.; Yu, J.; Klimas, A.; Williams, J.C.; Trayanova, N.A.; Entcheva, E. OptoGap is an optogenetics-enabled assay for quantification of cell–cell coupling in multicellular cardiac tissue. Sci. Rep. 2021, 11, 9310. doi:10.1038/s41598-021-88573-1.
18. Oliver-Gelabert, A.; García-Mendívil, L.; Vallejo-Gil, J.M.; Fresneda-Roldán, P.C.; Andelová, K.; Fañanás-Mastral, J.; Vázquez-Sancho, M.; Matamala-Adell, M.; Sorribas-Berjón, F.; Ballester-Cuenca, C.; et al. Automatic quantification of cardiomyocyte dimensions and connexin 43 lateralization in fluorescence images. Biomolecules 2020, 10, 1334. doi:10.3390/biom10091334.
19. Deng, R.; Cui, C.; Remedios, L.W.; Bao, S.; Womick, R.M.; Chiron, S.; Li, J.; Roland, J.T.; Lau, K.S.; Liu, Q.; et al. Cross-scale multi-instance learning for pathological image diagnosis. Med. Image Anal. 2024, 94, 103124. doi:10.1016/j.media.2024.103124.
20. Kamnitsas, K.; Ledig, C.; Newcombe, V.F.J.; Simpson, J.P.; Kane, A.D.; Menon, D.K.; Rueckert, D.; Glocker, B. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. Med. Image Anal. 2017, 36, 61–78. doi:10.1016/j.media.2016.10.004.
21. Deng, J.; Dong, W.; Socher, R.; Li, L.-J.; Li, K.; Fei-Fei, L. ImageNet: A large-scale hierarchical image database. Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR) 2009, 248–255. doi:10.1109/CVPR.2009.5206848.
22. Shorten, C.; Khoshgoftaar, T.M. A survey on image data augmentation for deep learning. J. Big Data 2019, 6, 60. doi:10.1186/s40537-019-0197-0.
23. Tan, M.; Le, Q.V. EfficientNetV2: Smaller models and faster training. Proc. 38th Int. Conf. Mach. Learn. (ICML), PMLR 2021, 139, 10096–10106.
24. Cui, Y.; Jia, M.; Lin, T.-Y.; Song, Y.; Belongie, S. Class-balanced loss based on effective number of samples. Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR) 2019, 9268–9277. doi:10.1109/CVPR.2019.00949.
25. Tajbakhsh, N.; Shin, J.Y.; Gurudu, S.R.; Hurst, R.T.; Kendall, C.B.; Gotway, M.B.; Liang, J. Convolutional neural networks for medical image analysis: full training or fine tuning?. IEEE Trans. Med. Imaging 2016, 35, 1299–1312. doi:10.1109/TMI.2016.2535302.
26. Yosinski, J.; Clune, J.; Bengio, Y.; Lipson, H. How transferable are features in deep neural networks?. Adv. Neural Inf. Process. Syst. (NeurIPS) 2014, 27, 3320–3328.
27. Akiba, T.; Sano, S.; Yanase, T.; Ohta, T.; Koyama, M. Optuna: A next-generation hyperparameter optimization framework. Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discov. Data Min. 2019, 2623–2631. doi:10.1145/3292500.3330701.
28. Bergstra, J.; Bardenet, R.; Bengio, Y.; Kégl, B. Algorithms for hyper-parameter optimization. Adv. Neural Inf. Process. Syst. (NeurIPS) 2011, 24, 2546–2554.
29. Saito, T.; Rehmsmeier, M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLoS ONE 2015, 10, e0118432. doi:10.1371/journal.pone.0118432.
30. Nadaraya, E.A. On estimating regression. Theory Probab. Appl. 1964, 9, 141–142. doi:10.1137/1109020.
31. Watson, G.S. Smooth regression analysis. Sankhyā Ser. A 1964, 26, 359–372.
32. Bensley, J.G.; De Matteo, R.; Harding, R.; Black, M.J. Three-dimensional direct measurement of cardiomyocyte volume, nuclearity, and ploidy in thick histological sections. Sci. Rep. 2016, 6, 23756. doi:10.1038/srep23756.