DOI:10.11918/j.issn.0367-6234.201711029

利用关键点检测算法的超声图像定点测量

朱 锴[1,2], 陶 攀[1,2], 付忠良[1], 陈晓清[1,2]

(1.中国科学院 成都计算机应用研究所, 成都 610041;2. 中国科学院大学,北京 100049)

摘 要: 超声图像的定点距离测量在临床医学上十分重要. 由于超声图像的噪声大,边缘模糊,因此关键点自动定位在超声图像中很有挑战性. 目前的关键点检测算法通常是针对单个关键点位置进行优化,难以在保证每个关键点检测精度的情况下得到准确的测量距离. 为使超声图像中关键点的精度和两个关键点之间的距离更加精确,本文提出一种基于级联卷积神经网络的关键点检测算法,该方法采用两个卷积网络从粗略到精细的对关键点进行定位. 首先利用第一个网络回归两个关键点的粗略位置,并将包含这两个关键点的小区域送入第二个网络. 然后本文提出一种加入距离修正的损失函数,作为第二个网络的优化目标,在第一个网络输出结果的基础上定位最终的关键点位置. 实验结果表明,本文提出的级联方法无论是相比传统的级联方式还是回归树方法,本文算法在超声图像的关键点定位上更为精准,并且在最终的距离测量精度上也有很大的提高,在评价标准下比传统级联方法检测精度上提升将近 30%.

关键词: 超声心动图;关键点定位;卷积神经网络;损失函数;自动测量

中图分类号: TP391.41 **文献标志码**: A **文章编号**: 0367-6234(2018)11-0067-07

Ultrasound image fixed point measurement based on landmark detection method

ZHU Kai[1,2], TAO Pan[1,2], FU Zhongliang[1], CHEN Xiaoqing[1,2]

(1. Chengdu Institute of Computer Applications, Chinese Academy of Sciences, Chengdu,610041, China;
2. University of Chinese Academy of Science, Beijing 100049, China)

Abstract: The measurement by fixed point of ultrasound image is very important in clinical medicine. Because of the large speckle noise and blurred edges in ultrasound images, landmark detection in echocardiography is quite challenging. Meanwhile, current landmark detection algorithm optimizes a single landmark position, and it's difficult to get accurate distance with guaranteeing the accuracy of each landmark. To get more accuracy result of landmark and the distance between two landmarks in ultrasound, cascaded convolution neural network is proposed to automatic detection landmark in echocardiography, our framework adopts two stages of carefully designed deep convolution networks that predict landmark location in a coarse-to-fine manner. Firstly, networks at the first level estimate positions of two landmarks coarsely, the patch includes two landmarks as input of the second stage. Then a loss function with distance correction term is proposed to optimize the second network, and gets the final landmark position based on the output of first stage. Experimental results show that compared to traditional network and regression tree, the proposed method could not only guarantee the accuracy of landmark position but also increase the accuracy of distance, compared with traditional cascaded convolution neural network, and the accuracy of distance is nearly 30% higher than the traditional method.

Keywords: echocardiography; landmark detection; convolution neural network; loss function; automatic measurement

超声心动图像是评估心脏功能获取心脏生理参数的重要手段[1],而二尖瓣的生理参数又是判断心脏健康与否的重要指标. 二尖瓣环内径的大小是判断二尖瓣是否正常的重要生理参数,通常医生获取二尖瓣环内径的大小是凭借自身经验标记两个瓣跟的关键点进行手动测量,其中二尖瓣根是位于二尖瓣膜和心室壁连接处有较为明显的转角特征,这样医生在术中操作非常不便. 由于心脏超声图像的形状、姿态和尺寸等变化多样,目前为止,并没有针对二尖瓣环内径关键点进行自动标注的有效算法,因此提出有效可行的二尖瓣环内径自动测量方法能够帮助医生高效的进行疾病诊断.

关键点检测算法在医学图像中一直有广泛应用,通过定位生理结构的一些关键位置能够对自动分析医学图像有很大的帮助. 在心脏图像中,很多

收稿日期: 2017-11-08
基金项目: 四川省科技厅重点研发项目(2017SZ0010); 四川省科技支撑计划项目(2016JZ0035)
作者简介: 朱 锴(1991—),男,博士研究生;
付忠良(1967—),男,研究员,博士生导师
通信作者: 付忠良,fzliang@ casit.com.cn

情况下进行关键点定位是进行图像分析的必要条件,如医学重要切面识别过程中运用关键点能够更为精确的定位生理学结构[2],心室分割过程中可以运用关键点检测算法辅助获取确切的边缘计算心室容积[3]. 关键点检测方法大体可以分为两类,一类是基于形状先验的轮廓模型方法,该类方法通常检测的关键点处于某个形状轮廓上;第二类为直接通过分类回归等思想定位目标关键点的方法,该类方法对关键的位置要求较低,但准确度不是很高. 第一类方法中,最为著名的为活动形状模型[4]和活动外观模型[5]方法,不少医学图像研究者采用这类方法进行图像分析,Yang 等[6]采用基于活动形状模型方法对超声图像进行分割,判断颈动脉粥样硬化性疾病,Vargas 等[7]使用了活动外观模型对超声心动图的左心室进行分割. 这类基于形状模型的方法通常根据训练数据得到平均形状模型,能够对关键点有全局的形状约束,在提取局部特征的情况下加入了全局形状外观约束. 但是这种方法通常需要有良好的初始轮廓位置,并且提取的特征一般是人工设计有局限性的特征. 在不通过形状约束的定位关键点方法里,有从事医学图像分析的研究者采用随机森林结合 HOG 特征检测右心室的关键点[8],为让检测结果能够更好的适应解剖学结构姿态、形状的变化,学者 Ozan 等[9]在决策森林回归方法的基础上引入了分层思想让检测结果更为精确. 近年来,随着深度学习的大力发展,将神经网络运用于解剖学结构关键点检测的研究人员也越来越多,Dong 等[10]在 3D 的医学图像上使用卷积神经网络(CNN)对股骨关键点进行定位,并结合图割方法分割股骨表面,Florin 等[11]使用 CNN 对超声和 CT 图像进行解剖学关键点的检测. 这类方法通常能够得到更好的检测精度,但是依赖于大量的训练数据. 虽然上述关键点检测在医学图像上取得了一定的成绩,但实验中发现这些方法直接应用于二维超声图像二尖瓣环内径的测量仍然存在一些问题:

1)测量二尖瓣环内径仅仅需要标记两个关键点,形状模型方法需要标注多个关键点才能利用其形状轮廓进行约束,并且优化目标是整个轮廓,关心的是所有关键点的平衡,并不是实际需要测量的两个关键点的距离.

2)超声心动图的边界模糊,采用回归思想的方法虽然能够大致获取关键点的位置,但是现有算法对每个关键点检测都有一定误差,这样对测量距离的稳定性有很大影响.

为解决上述问题,本文提出了一种基于 CNN 的回归模型来对二尖瓣环的两个关键点进行定位,CNN 最初由 LeCun 等[12]在 1998 年提出的,再由 2012 年 Hiton 等[13]对其改进将算法性能进行大幅度提高,该结构使用网络自己学习特征,避免了人工提取特征的局限性. 现有很多关键点检测算法都是基于 CNN 的,Deng 等[14]提出了新的数据增广策略采用了单一的网络结构对人脸关键点进行定位. 为了让定位更为精准,许多学者采用级联 CNN 的方式对关键点由粗糙到精细的定位. Sun 等[15]在 2013 年采用级联的 CNN 对人脸关键点进行定位,后又引入难样例鉴别机制结合级联的 CNN 进行服装关键点检测[16],之后 Zhang 等[17]采用了三个网络对人脸关键点有粗到细的定位. 由于此类算法采用的是直接回归关键点的位置,损失函数使用的都是欧式距离损失,在预测关键点的时候通常是孤立的预测每一个关键点,将其直接运用于测量二尖瓣环内径得到的效果并不理想.

本文提出的基于 CNN 的二尖瓣环内径测量方法,通过使用 CNN 回归关键点的位置,在欧式距离损失函数基础上加入了距离损失函数,在关键点定位的同时也考虑了距离准确度的问题,实验证明,该方法能够在大部分图像上得到令医生满意的结果.

1 基于级联 CNN 的关键点检测算法

基于 CNN 的关键点检测方法是由 Sun 在人脸上率先使用的[15],之后大量算法在都是在此基础上改进并对每个关键点由粗到细的进行回归预测[16-17]. 该算法采用三层级联的卷积网络进行:第一级的卷积网络分为三个单独的卷积网络,分别回归出不同部位的关键点粗略的位置,第二级和第三级的卷积网络以前一级的预测点为中心的小块(patch)作为输入,对前一级的网络输出做微调. 每个网络框架见图 1. 其中每个卷积层有多个特征矩阵组成,运用于整个特征图像,减少了参数数目,目的是提取图像特征,然后通过池化层降采样,之后通过全连接层输出所需特征点坐标. 其中卷积层的输入输出可用式(1)来表达

$$Y_1 = \tanh(W \otimes Y + B). \tag{1}$$

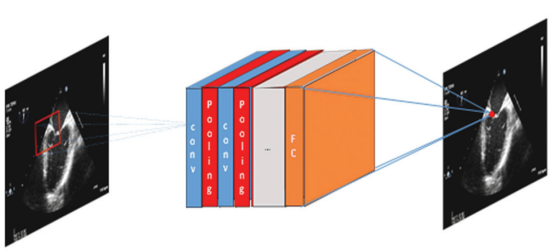


图 1 卷积神经网络回归关键点框架

Fig.1 Pipeline of landmark detection with convolution neural network

式中：Y 代表输入特征图，Y_1 代表输出特征图，W 为权重参数，B 是偏置项，文中采用的是双曲正切函数(tanh)作为激活函数，损失函数采用的是式(2)表达的欧式距离损失(Euclidean loss，EL)，即方根均差(root mean squared error，RMSE)．

$$R_{\mathrm{RMSE}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}. \qquad (2)$$

式中：y_i 为 n 个关键点构成的向量$(x_1, y_1, x_2, y_2, \cdots, x_n, y_n)$ 第 i 维的数值，$\hat{y}_i$ 为预测向量第 i 维的数值．

2　改进的基于级联 CNN 的关键点检测算法

2.1　级联结构的改进

本文提出的超声图像关键点检测算法由两个 CNN 级联构成，目的是检测二尖瓣根部的两个关键点并计算其长度．传统检测关键点的 CNN 级联方式见图 2(a)，先通过第一个网络定位目标关键点的粗略位置，第二级网络利用第一级预测点为中心的小区域作为输入进行微调．采用传统方法得到的关键点之间关联性较差，预测距离差距较大．为解决这个问题，本文的级联方式见图 2(b)，和传统级联方法相同，第一阶段是对关键点位置做粗略定位，采用的是欧式距离损失函数作为优化目标，第二阶段网络在第一阶段预测结果的基础上，选取包含两个预测点的小区域作为输入，并采用带有距离修正的欧式损失函数作为优化目标，得到最终的预测结果．第一个网络结构见图 2(c)，输入图像尺寸较大，本文采用含有 6 个卷积层和 2 个全连接层的网络进行回归，获得关键点的大概位置，第二个网络见图 2(d)，本文选用的是 5 个卷积层和 2 个全连接层的结构，回归得到最终的结果．

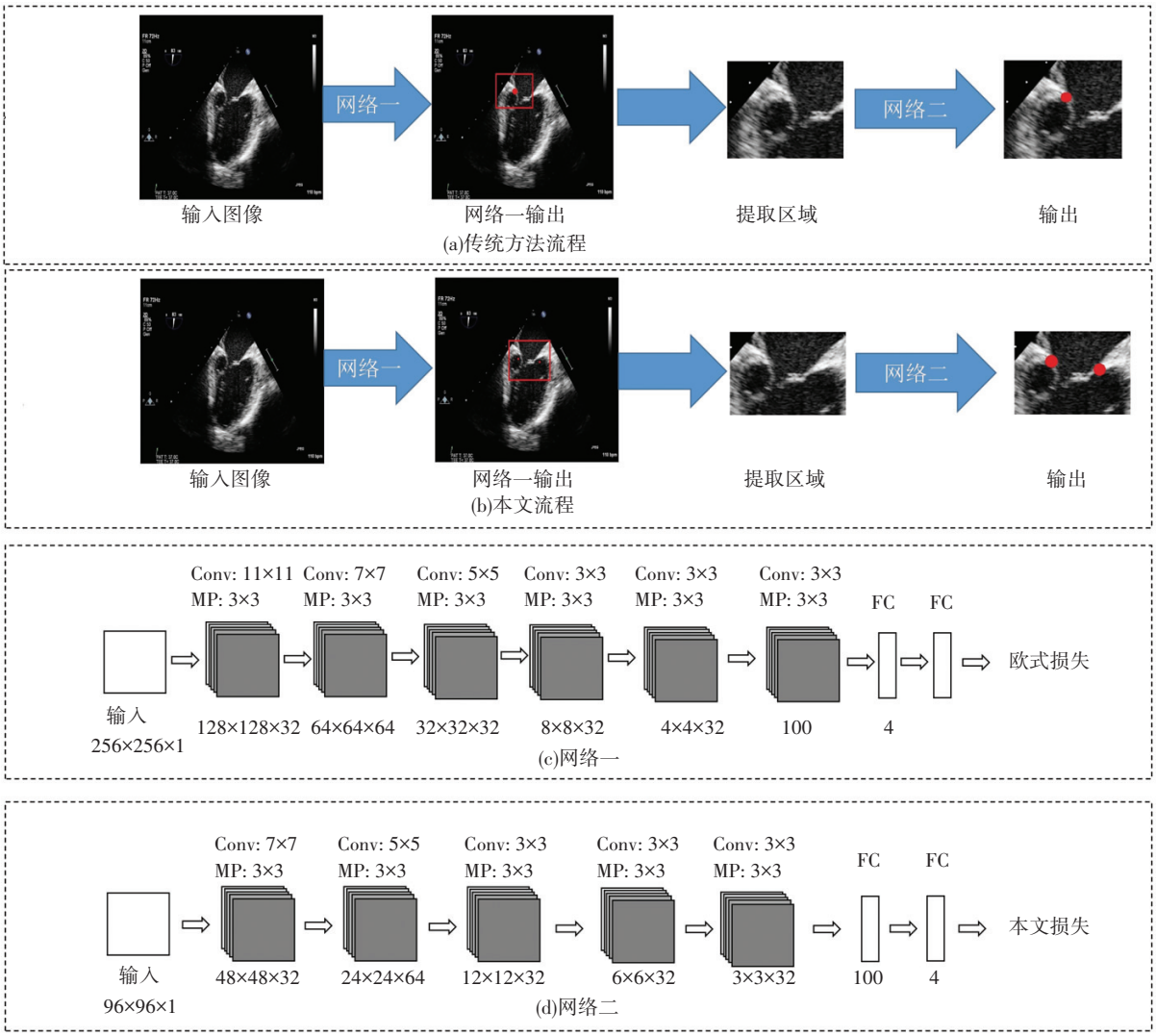


图 2　级联流程及网络结构

Fig.2　Pipeline of method and architectures of two networks

2.2 损失函数的改进

目前大部分基于回归的关键点检测算法所采用的损失函数都是欧式距离损失,这种损失函数在关键点回归过程中表现良好,针对每一个关键点检测结果都有较高的精度. 但在距离测量中,该损失函数并不能体现距离这一重要因素,仅考虑孤立的每个关键点作为损失标准在距离自动测量中具有很大的局限性.

如图 3 所示, 绿色的点 A 和 B 代表真实的关键点位置,A_1 代表 A 点的预测位置,B_1 和 B_2 代表两种不同的 B 点的预测位置,对于欧式距离损失来说,B_1 和 B_2 带来的损失值是一样的,显然,预测结果 1 和 2 相对真实距离 AB 是有很大的不同的,预测结果 2 更为准确,所以实际上更希望出现的是 A_1B_2 这种情况,因为这样预测结果更为精确. 所以针对只用检测两个关键点测距的目标,本文在原有欧式损失的基础上引入距离损失的度量, 提出一种新的带距离修正的损失函数(loss function with distance correction term, DL). 如式(3)所表示的. 其中 $(x_1^{\mathrm{T}},y_1^{\mathrm{T}})$ 和 (x_2^{T},y_2^{T}) 代表两个真实关键点的坐标, $(x_1^{\mathrm{P}},y_1^{\mathrm{P}})$ 和 $(x_2^{\mathrm{P}},y_2^{\mathrm{P}})$ 代表预测的两个关键点坐标, L_1 和 L_2 分别代表每个关键点的预测值和真实值的距离,L_T 和 L_P 分别表示的是两个真实关键点距离和预测的关键点距离.

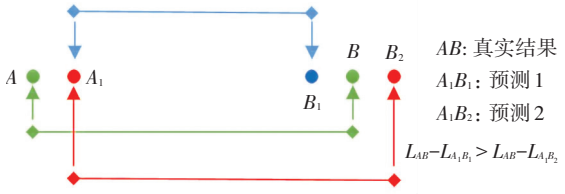


图 3 距离测量示意

Fig.3 Diagram of measuring distance

$$D_{\mathrm{disloss}}=\sqrt{\frac{1}{2}(\mathrm{L}_1{}^2+L_2{}^2)+\lambda\left|L_{\mathrm{T}}{}^2-L_{\mathrm{P}}{}^2\right|},$$

$$L_1=\sqrt{(x_1^{\mathrm{T}}-x_1^{\mathrm{P}})^2+(y_1^{\mathrm{T}}-y_1^{\mathrm{P}})^2},$$

$$L_2=\sqrt{(x_2^{\mathrm{T}}-x_2^{\mathrm{P}})^2+(y_2^{\mathrm{T}}-y_2^{\mathrm{P}})^2},$$

$$L_{\mathrm{T}}=\sqrt{(x_1^{\mathrm{T}}-x_2^{\mathrm{T}})^2+(y_1^{\mathrm{T}}-y_2^{\mathrm{T}})^2},$$

$$L_{\mathrm{P}}=\sqrt{(x_1^{\mathrm{P}}-x_2^{\mathrm{P}})^2+(y_1^{\mathrm{P}}-y_2^{\mathrm{P}})^2}. \tag{3}$$

3 实验及分析

3.1 实验数据

实验数据来源四川华西医院麻醉科,获取左室二腔心切面视频 135 个,其中 100 个视频用于训练图片抽取,其余 35 个视频用于测试图片抽取. 每个视频根据视频长度选取 3~5 帧标注,共获得训练数据图片 387 张,测试数据图片 122 张. 每幅图像标注两个关键点,图像大小为 800 * 600 pixel,全部处理为灰度图像. 对于第一个网络,对训练集图像执行如下数据集扩充方法:

1)对原始图像矩阵 X 做顺时针旋转操作,即 $X=X\cdot R$,其中 R 见式(4), 旋转后没有填充的部分使用图像的平均值填充,旋转角度随机选取 5°~15°,得到 774 张图片.

$$R=\begin{bmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{bmatrix}. \tag{4}$$

2)对 1 的结果做对比度变换,即 $X=(\delta*X)+(1-\delta)*X_{\mathrm{mean}}$,其中 X_{mean} 代表所有像素平均值,δ 为对比度参数,本文将从随机 0.6 ~ 0.8 中随机选取. 通过这一步能够得到 1 548 张训练图像.

对于第二个级联网络,文章从每个图像中选取包含两个关键点的小区域作为输入, 获取区域公式如下表达. $(x_{\text{left-top}},y_{\text{left-top}})(x_{\text{right-top}},y_{\text{right-bottom}})$ 代表新生成区域左上角和右下角坐标,(x_1,y_1) 和 (x_2,y_2) 分别代表标注两个关键点位置. R_{x1}, R_{y1}, R_{x2}, R_{y2} 在试验中是从 15 ~ 80 之间生成的随机数,表达式如式(5) 所示.

$$\begin{aligned}
&x_{\text{left-top}}=X_{\min}-R_{\mathrm{x1}},X_{\min}=\min(x_1,x_2);\\
&y_{\text{left-top}}=Y_{\min}-R_{\mathrm{y1}},Y_{\min}=\min(y_1,y_2);\\
&x_{\text{right-bottom}}=X_{\max}+R_{\mathrm{x2}},X_{\max}=\max(x_1,x_2);\\
&y_{\text{right-bottom}}=Y_{\max}+R_{\mathrm{y2}},Y_{\max}=\max(y_1,y_2).
\end{aligned} \tag{5}$$

最后从 387 幅训练图像中每幅选取 15 个小区域作为训练数据. 两个网络的部分数据见图 4.

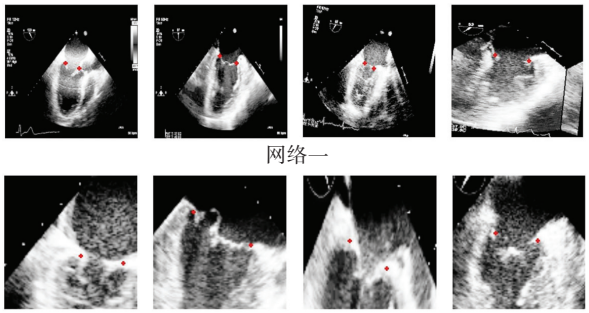


图 4 训练样本展示

Fig.4 Examples of train dataset

3.2 评价标准

本文采用式(2)展示的方均根差对所有关键点检测测试样本进行评估. 并且定义误差不超过图像宽度 $e<5\%$ 的检测点为正确检测点. 距离误差公式见式(6)

$$e=\frac{\sqrt{(x_{\mathrm{p}}-x_{\mathrm{t}})^2+(y_{\mathrm{p}}-y_{\mathrm{t}})^2}}{w}. \tag{6}$$

式中:(x_p,y_p) 和 (x_t,y_t) 分别代表预测的关键点位

置和标注的关键点位置，w 代表图像宽度. 对于测量距离，本文将两点之间距离检测误差小于 $e_{\text{dis}} < 15\%$ 的距离视为正确距离，这是医生可接受的误差范围. 对测量的距离误差衡量表达见式(7)，其中 l_t 为真实距离，l_p 为预测距离.

$$e_{\text{dis}} = \frac{|l_{\text{p}} - l_{\text{t}}|}{l_{\text{t}}}. \tag{7}$$

将 $e_{\text{dis}} < 15\%$ 并且 $e < 5\%$ 的检测结果视为正确结果，定义检测成功率 T_a 见式(8)，其中 f_T 为检测成功的样本数目，N 为测试总样本数目.

$$T_a = \frac{f_T}{N} \times 100\%. \tag{8}$$

3.3　实验结果及分析

实验环境：操作系统为 ubantu14.04，显卡为 GTX titan x，代码是采用基于 caffe 的 python 实现的.

实验结果见表 1，展示在采用不同损失函数对最终二尖瓣距离测量精度的影响，并同文献[15]的级联方式以及文献[18]采用回归树的方法进行比较. 即在第一级级联状况相同的情况下，将第二级改成普通的针对单个关键点进行精细化调整. 从表中结果前两行可以看出仅采用第一层网络结构对点的定位效果很差，对左侧关键点和右侧关键点检测成功率均很低，最终的距离测量结果不理想，并且两种损失函数对最终距离测量准确度影响并不大. 在采用两层网络的情况下，两个关键点定位效果明显好转，关键点检测成功率均达到了 90% 以上，对最终距离的准确度也大大提高. 但从表中网络二选用不同的损失函数可以看出，在第一个网络采用欧式损失的情况下，第二个网络采用本文提出的损失函数在关键点定位精度上和采用普通欧式损失的网络区别并不大，但是在最终的距离准确度上效果要好很多，获得的平均距离误差也更小. 而两个网络都采用本文提出的损失并不能提升定位效果和测量精度. 并且对比文献[15]的级联方法，相对普通针对每个关键点进行精细回归的级联方法，在关键点检测和距离测量的结果上，本文提出方法精度上提升很大. 在和传统学习算法的对比上，将在第一个网络输出的基础上，采用文献[18]的回归算法对检测结果进行调整，结果也显示本文的算法远远优于传统回归算法.

表 1　不同级联方式和损失函数实验对比

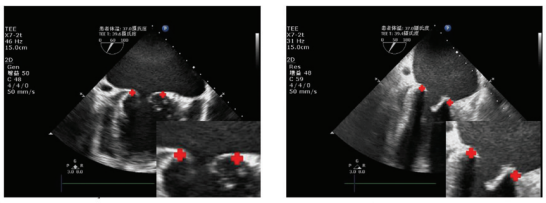
Tab.1　Result of different structures and loss function

算法	损失函数	左关键点	右关键点	RMSE	检测成功率	距离平均误差
网络一	EL	59.8%	70.5%	12.18	41.8%	12.3%
	DL	64.8%	68.0%	12.52	39.3%	14.0%
网络二	EL+EL	92.6%	96.0%	7.30	82.7%	9.3%
	EL+DL	97.5%	93.4%	7.04	86.1%	7.5%
	DL+DL	96.7%	94.3%	7.46	84.4%	8.1%
文献[15]	EL	82.7%	92.6%	8.95	65.6%	9.4%
文献[18]	—	51.6%	46.7%	22.88	33.6%	13.0%

不同算法下的预测效果见图 5. 图 5(a)展示的是医生标注的结果，图 5(b)展示的在只使用一个网络并且损失函数采用欧式距离损失下得到的结果，很明显的可以看到仅仅采用一个网络得到的结果并不太理想，图 5(c)是两个级联网络都使用欧式损失得到的结果，图 5(d)是将第二个网络采用本文提出的损失函数得到的结果，在采用本文提出的损失函数后再检测二尖瓣环内径的距离上能够得到更为准确的结果，图 5(e)为采用文献[15]的级联方式得到的实验结果，图 5(f)为采用文献[18]的回归树算法得到的实验结果. 无论是从视觉上还是数值上，本文提出的算法都能够达到较为准确的结果. 并且算法在定位关键点和预测距离上做一个较好的平衡.

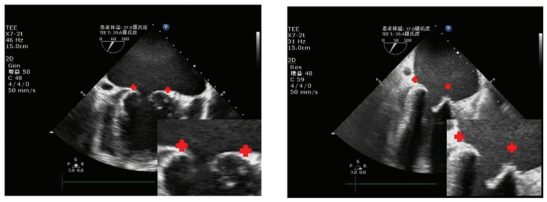
训练损失和迭代轮次的关系见图 6. 训练过程中采用固定学习率为 0.001 方式，采用带动量的优化方法. 在本文数据量不大的情况下，算法能快速收敛.

在第一级网络选取相同的情况下，第二个网络损失函数采用不同的 λ 对最终模型表现影响见图 7 所示. 从图中可看出，在参数 λ 增大的情况下，对关键点的定位精度和最终的距离精度有一定的影响，但是加入距离修正项的损失函数能够让距离检测更好，实验中可以看出在 $\lambda = 1$ 的情况下能够对点的检测精度 和距离测量的准确度有较好的表现. 在 λ 选取过小的情况下，距离约束不够明显，导致最终测量结果不理想，而 λ 过大的情况下对关键点检测精度会有 较大影响，也会导致距离测量结果出现较大误差.



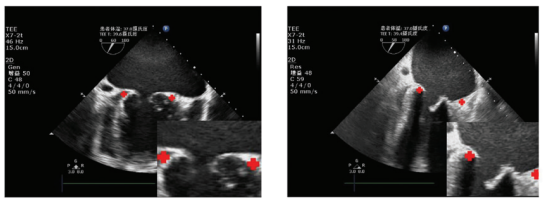
l=3.26 cm *l*=3.28 cm

(a)标注结果



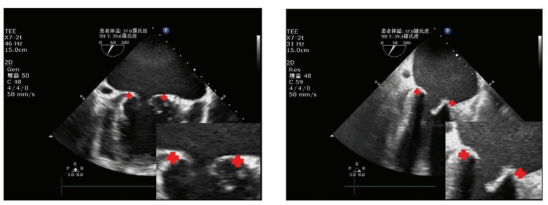
l=3.52 cm *l*=3.47 cm

(b)网络一



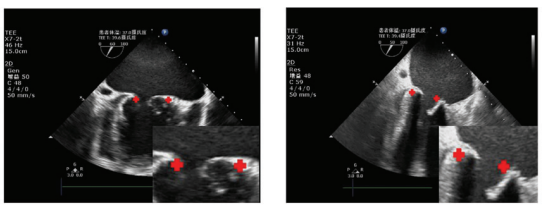
l=4.97 cm *l*=4.60 cm

(c)EL+EL



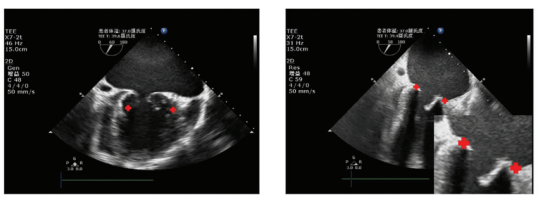
l=3.50 cm *l*=3.48 cm

(d)本文算法



l=3.30 cm *l*=2.60 cm

(e)文献[15]



l=3.91 cm *l*=3.36 cm

(f)文献[18]

图 5 不同级联方式和损失函数检测样例

Fig.5 Examples of different structures and loss function

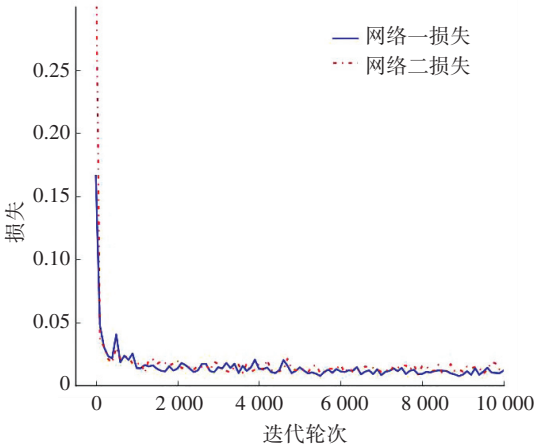
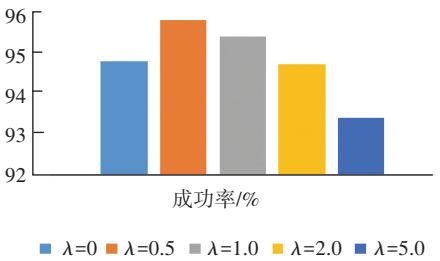
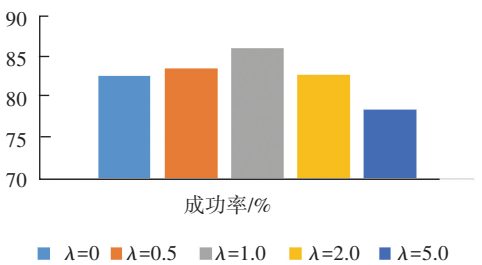


图 6 两个阶段网络训练损失和轮次的关系

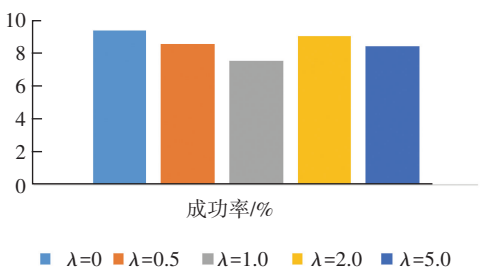
Fig.6 Losses across epoch of two stages



(a) 不同λ下的关键点检测精度



(b) 不同λ下的距离检测精度



(c) 不同λ下距离检测的平均误差

图 7 参数 λ 对测量结果的影响对比

Fig.7 Results of different λ

4 结 论

本文提出一种基于关键点检测的超声图像的二尖瓣环内径测量方法,该方法通过两个 CNN 级联对二尖瓣环内径的两个关键点进行了有粗糙到精细的定位. 另外,通过加入带距离修正的损失函数作为

优化目标，相对仅仅关心关键点精度的欧式距离损失，该损失让最终检测的距离更为精准. 并且本文探讨了距离损失权重参数 λ 对最终结果的影响. 本文是首先将关键点检测算法运用于超声图像二尖瓣环内径的自动测量的，通过实验证明，在误差范围内算法的测量结果更接近实际距离.

参考文献

[1] HEIDENREICH P A, MADDOX T M, NATH J. Measuring the quality of echocardiography using the predictive value of the left ventricular ejection fraction[J]. Journal of the American Society of Echocardiography, 2013, 26(3): 237. DOI: 10.1016/ j. echo. 2012.12.011

[2] LU X, JOLLY M P, GEORGESCU B, et al. Automatic view planning for cardiac MRI acquisition[J]. Medical Image Computing and Computer-Assisted Intervention-MICCAI 2011, 2011: 479. DOI: 10.1007/978-3-642-23626-6_59

[3] 纪祥虎，高思聪，黄志标，等. 基于 Centripetal Catmull-Rom 曲线的经食道超声心动图左心室分割方法[J]. 四川大学学报：工程科学版，2016，48(5)：84. DOI:10.15961/j.jsuese.2016.05.013

JI X, GAO S, HUANG Z, et al. Left ventricle segmentation in transesophageal echocardiography based on centripetal catmull-rom curve[J]. Sichuan Daxue Xuebao (Gongcheng Kexue Ban)/Journal of Sichuan University (Engineering Science Edition), 2016. DOI: 10.15961/j.jsuese.2016.05.013

[4] COOTES T F, TAYLOR C J, COOPER D H, et al. Active shape models-their training and application[J]. Computer Vision and Image Understanding, 1995, 61(1): 38. DOI:10.1006 /cviu.1995.1004

[5] COOTES T F, EDWARDS G J, TAYLOR C J. Active appearance models[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23(6): 681

[6] YANG X, JIN J, XU M, et al. Ultrasound common carotid artery segmentation based on active shape model.[J]. Computational and Mathematical Methods in Medicine, 2013,(2013-3-6), 2013, 2013(2):345968.DOI: 10.1155/2013/ 345968

[7] VARGAS-QUINTERO L, ESCALANTE-RAMÍREZ B, MARÍN L C, et al. Left ventricle segmentation in fetal echocardiography using a multi-texture active appearance model based on the steered Hermite transform[J]. Computer Methods and Programs in Biomedicine, 2016, 137: 231.DOI: 10.1016/j.cmpb. 2016. 09.021

[8] SEDAI S, ROY P K, GARNAVI R. Right ventricle landmark detection using multiscale HOG and random forest classifier[C]//Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on. New York: IEEE, 2015: 814. DOI: 10. 1109/isbi. 2015. 7163996

[9] OKTAY O, BAI W, GUERRERO R, et al. Stratified decision forests for accurate anatomical landmark localization in cardiac images [J]. IEEE Transactions on Medical Imaging, 2017, 36(1): 332. DOI: 10.1109/tmi.2016.2597270

[10] YANG D, ZHANG S, YAN Z, et al. Automated anatomical landmark detection ondistal femur surface using convolutional neural network[C]//Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on. New York: IEEE, 2015: 17. DOI: isbi. 2015.7163806

[11] GHESU F C, GEORGESCU B, MANSI T, et al. An artificial agent for anatomical landmark detection in medical images[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Athens: Springer International Publishing, 2016: 229. DOI: 10.1007/978-3-319-46726-9_27

[12] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278. DOI: 10.1109/5.726791

[13] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[C] //Advances in Neural Information Processing Systems. New York MIT Press Curran Associates Inc.2012: 1097. DOI: 10.1145 /3065386

[14] DENG Z, LI K, ZHAO Q, et al. Effective face landmark localization via single deep network [J]. ArXiv Preprint ArXiv: 1702. 02719, 2017

[15] SUN Y, WANG X, TANG X. Deep convolutional network cascade for facial point detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Portland: IEEE Computer Society. 2013: 3476. DOI: 10.1109/cvpr.2013. 446

[16] LIU Z, YAN S, LUO P, et al. Fashion landmark detection in the wild[C]//European Conference on Computer Vision. Amsterdam: Springer International Publishing, 2016: 229. DOI: 10.1007 /978-3-319-46475-6_15

[17] ZHANG K, ZHANG Z, LI Z, et al. Joint face detection and alignment using multitask cascaded convolutional networks[J]. IEEE Signal Processing Letters, 2016, 23(10): 1499. DOI: 10.1109/lsp.2016.2603342

[18] KAZEMI V, SULLIVAN J. One millisecond face alignment with an ensemble of regression trees[C]// Computer Vision and Pattern Recognition. Columbus: IEEE, 2014:1867. DOI: 10.1109 /cvpr. 2014.241

（编辑　苗秀芝）