Cardiac Image Analysis with Deep Learning Methods

Aliasghar Mortazi\textsuperscript{1}, G.Z. Papadakis\textsuperscript{2}, U. Teomete\textsuperscript{3}, Ulas Bagci\textsuperscript{1}

\textsuperscript{1}Center for Research in Computer Vision, University of Central Florida, Orlando, FL., US.

\textsuperscript{2}Computational Bio-Medicine Laboratory (CBML), Heraklion, Crete, Greece.

\textsuperscript{3}University of Miami Hospital.
Outline

- Evolution in deep learning and CNN designs
- Deep learning in Left Ventricle analysis
- Deep learning in Myocardium of Left Ventricle analysis
- Deep learning in Right Ventricle analysis
- Left Atrium and Pulmonary Veins analysis with deep learning
- Coronary Arteries and Calcium Scoring with Deep Learning
- Multi substructure segmentation with Deep Learning
- Challenges in Cardiac Image Analysis with Deep Learning
- Effectiveness of Deep Learning in Future of Cardiac Image Analysis
Evolution in deep learning and CNN designs

- Evolution in Deep Learning Architectures & Connections
- Evolution in Optimization Methods in CNNs
- Deep Reinforcement Learning
- Capsules
At the beginning, the data flow in the CNNs were layer-to-layer and sequentially and there were no connections among different layers.

The classification layer was applied only to high-level features.

The ResNet [3] (for recognition) and U-Net [2] (for segmentation) task were first architectures tried to use low-level features in classification layer by putting some connections from beginning layers to ending layers in network.

The encoder-decoder architecture proposed by Noh and et al. For segmentation. The data flow were only layer-by-layer [1].
In order to use all low-level features in the last layer and also to solve the gradient vanishing problem, all the layers were connected together in **densely connected CNN** [4] for detection task.

Then, the **Tiramisu** architecture [5] was proposed by combining the idea of U-Net and densely connected CNN (by adding connections from encoder and decoder) for segmentation task.
A RNN is a network with loop which allows the network to keep the information though the time.

LSTM is one of the famous RNNs solve long-term dependency problem

i.e in cine-cardiac images it can be used to learn the relationship among the slices in different times

GAN method was introduced by Goodfellow which two networks were used: Discriminator and Generator

Generator is training to generate images from noise to fool discriminator

Discriminator is training to discriminate the images generated by generator (fake images) from real images
The Optimization methods have been played a key role in making the deep learning practical by reducing the training time and increasing convergence speed [8]

<table>
<thead>
<tr>
<th>Optimization Method</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Batch Gradient descent</td>
<td>It is guaranteed to converge to the global minimum for convex function</td>
</tr>
<tr>
<td></td>
<td>Computationally it is expensive and need huge amount of memory</td>
</tr>
<tr>
<td>Stochastic Gradient descent</td>
<td>SGD's fluctuation enables it to jump to new and potentially better local</td>
</tr>
<tr>
<td></td>
<td>minimum</td>
</tr>
<tr>
<td></td>
<td>It complicates convergence to the exact minimum</td>
</tr>
<tr>
<td>Mini-Batch Gradient descent</td>
<td>Typically the algorithm of choice when training a neural network</td>
</tr>
<tr>
<td></td>
<td>Reduces the variance of the parameter updates</td>
</tr>
<tr>
<td>Momentum SGD:</td>
<td>Momentum is a method that helps accelerate SGD in the relevant</td>
</tr>
<tr>
<td></td>
<td>Not intelligence enough! blindly following the slope!</td>
</tr>
<tr>
<td>Nesterov accelerated</td>
<td>Can speed up the convergence by providing more smarter updates</td>
</tr>
<tr>
<td></td>
<td>Can’t adapt the updates to each individual parameter in network</td>
</tr>
</tbody>
</table>

**Momentum-based optimizers**

**Gradient descent optimizers**
Adaptive optimizers were one of the major breakthroughs in deep learning, helping networks converge faster and achieve better accuracy [8].

- **Adagrad**
  - It adapts the learning rate to the parameters, performing smaller updates.
  - It causes learning rate eventually become infinitesimally small.

- **Adadelta**
  - Expansion of Adagrad to solve learning rate problem.
  - It is not updating momentum adaptively.

- **Adam**
  - Can update both learning rate and momentum adaptively.

- **Warm Restart Optimizers**
  - **SGD with restarts**
    - Learning rate is restart after some epochs to avoid local minimum.
    - It requires 2 to 4 times fewer epochs.
Reinforcement Learning (RL) is a learning method based on the reward/punishment which an agent can receive from environment for its performance.

A RL includes 6 components:
1) **Agent**: Who can interact with environment
2) **Action**: Set of possible moves for agent
3) **Environment**: Definition of what agent can do
4) **Policy**: the strategy that the agent employs
5) **Rewards**: The feedback by which we measure the success
6) **State**: Immediate situation in which the agent finds itself

A CNN can be seen as an agent that learns to map state-action pairs to rewards.
Capsules

- The CapsulNet was introduced with Sabour and et al. which made CNNs Equivariance. [11]

- One of major problem of CNNs was lack of enough power to generalize from training data to rotated data (non-Equivariance).

- **SegCaps** was proposed in [12] and it was an extension of CapsulNet for doing segmentation in medical imaging.
Deep learning in Left Ventricle analysis

- Automatic Localization of the Left Ventricle in MRI
- LV Segmentation on 3D Echocardiography using CNN
- The Segmentation of the LV of From Ultrasound
- Tracking the LV of the Heart from Ultrasound Data
- Recognizing End-Diastole and End-Systole
1- Purposes:
• Automatic localization of the left ventricle (LV) in cardiac MRI images with CNN-based methods
• Useful for Automatic segmentation, functional analysis, and content based retrieval of images

2- Methods
• A network with seven layers including convolutional layers following by fully connected layers are used for this task
• A pyramid-of-scales is used to account for the variations of heart size during the localization
• A given input image is resized to four different scales.

3- Experiments & Results
The accuracy of 98% and sensitivity of 84% were achieved by proposed method. Some of the qualitative results are illustrated.
1- Purpose:
• Segmentation of LV endocardium from 3D echocardiography
• It can provide some clinical indices such as ventricular volume, and ejection fraction and also can be used for the analysis of anatomic structure of ventricle
• Full-automatic method by combining the deep learning and deformable model

2- Methods
• A CNN is used for localization of LV and then a stacked autoencoder is used to refine the output of the CNN. Finally, a GVF-snake is applied for have final LV segmentation as is illustrated in below figure:

3- Experiments & Results
• The 3DE data from CETUS challenge 2014 are used. There are 45 sequences of 3D ultrasound volumes, 15 volumes of training set and 30 volumes of test dataset.
• The mean surface distance and hausdorff surface distance of 2.2 mm and 8.34 mm for ED, and 2.56 and 8.46 for ES were achieved.
1- Purpose:
- Automatic segmentation of the LV of the heart in ultrasound images
- Manual segmentation of the LV is a tedious and time demanding task and it is prone to poor repeatability

![Example of manual annotation](image)

2- Methods
- Deep belief networks (DBN) is used as a rigid classifier takes as input an image region, and the output is the probability that the region contains an LV.
- Then, a nonrigid classifier takes a profile line perpendicular to the LV contour and outputs the most likely location of the LV contour
- A rigid classifier is a pixel-based method designed for the task of object detection and recognition, and the nonrigid classifier is a pixel-based method designed for the task of segmentation

3- Experiments & Results
- Data set of diseased cases containing 400 annotated images (from 12 sequences) and another data set of normal cases comprising 80 annotated images (from two sequences), where both sets present long axis views of the LV

Proposed method: LV segmentation to an image is independent of the original cause (i.e., the imaging of the LV of the heart) [15]

DL approach: an generative model learns the LV image generation process and, then, a model is trained based on this generative model [15]
1- Purpose:
• Automatic tracking and segmentation of the LV from ultrasound images
• The automatic tracking will provide to a quantitative functional analysis of the heart, such as the ejection fraction estimation for clinician

2- Methods
• A new LV tracking algorithm based on Sequential Monte Carlo (SMC) methods
  • Transition model: makes use of the prior information that at each time instant, the heart is either expanding (diastole) or contracting (systole)
  • Observation model: is based on deep learning architectures, which involves a statistical pattern recognition model
• DNN is used to produce more abstract feature spaces for classification

3- Experiments & Results
• Average distance of 3.2 mm and Hausdorff distance of 19.1 mm were achieved by proposed method
• The yellow, solid line displays the manual annotation, while the magenta dashed line shows

Block diagram containing all steps of the tracking algorithm [16]
Recognizing End-Diastole and End-Systole[17]

1- Purposes:
• TempReg-Net is used to accurately identify end-systole and end-diastole from MRI sequences, by integrating the CNN with the RNN
• Accurate measurement of left ventricular volumes and Ejection Fraction from cine-MRI

A typical example of cardiac sequences[17]

2- Methods
• A CNN is used to encode the spatial pattern of the cardiac image sequence
• Then, a RNN is used to decode the temporal information of the extracted features from CNN
• The joint network can be trained to learn the complex spatial and temporal patterns of the cardiac sequences

An overview of the proposed framework[17]

3- Experiments & Results
• The dataset comprises of cardiac sequences and consists of around 113,000 cardiac frames are used for training and evaluation
• Average difference of 0.4 frames is achieved by proposed method for recognizing ED and ES
Deep learning in Myocardium of Left Ventricle analysis

- Direct Estimation of Regional Wall Thicknesses
- EF Classification in Transthoracic Echocardiography
- Analysis of the myocardium in coronary CT angiography
- Detection and Classification of Coronary Artery Plaque
Direct Estimation of Regional Wall Thicknesses

1- Purposes:
- Accurate and direct estimation of regional wall thicknesses (RWT) of left ventricular (LV) myocardium from cardiac MR sequences
- A dataset includes 2900 images of short-axis cine MR images paired with manually obtained ground truth values of RWT are used

Illustration of RWT for short-axis view cardiac MR image

2- Methods
Residual Recurrent Neural Network (ResRNN) consisting of a RNN path (for modeling the temporal and spatial dependencies) and CNN path (to obtain a preliminary estimation of RWT) is used for estimation RWT

Overall view of proposed method

3- Experiments & Results
Cardiac RWT with Mean Absolute Error of 1.44mm (less than 1-pixel error), when validated on cardiac MR sequences of 145 subjects, was achieved
EF Classification in Transthoracic Echocardiography [19]

1- Purposes:
• Measuring the ventricular ejection fraction (EF) through transthoracic echocardiography (TTE)
• EF is currently determined through a semi-automatic process that requires manual delineation of the left ventricle
• Ejection fraction in classified to four classes based on TTE exams

2- Methods:
• 3D-Convolutional Neural Network (3D-CNN) trained on a dataset constructed with exams from a cardiology reference center
• 3D-CNN model was used to integrate temporal knowledge from TTE cineloops in the model’s input by considering time the third dimension.

3- Experiments & Results:
• A dataset contains 8715 exams with 30 sequential frames were extracted to capture a cardiac cycle
• F1 score was computed for each class, with the following results: Class 1 - 71.3%, Class 2 - 63.3%, Class 3 - 72.3% and Class 4 - 54.6%

Existing and proposed pipelines to extract information from TTE exams (top and bottom rows, respectively) [19]
Analysis of the myo in coronary CT angiography [20]

1- Purposes:
- Fractional flow reserve (FFR) measurement, performed during invasive coronary angiography (ICA), is most often used in clinical practice.
- To reduce the number of ICA procedures, we present a method for automatic identification of patients with functionally significant coronary artery stenoses.

2- Methods:
- First, the LV myocardium is segmented using a multiscale convolutional neural network (CNN).
- To characterize the segmented LV myocardium, it is subsequently encoded using unsupervised convolutional autoencoder (CAE).
- The LV myocardium is divided into a number of spatially connected clusters.
- Then, patients are classified according to the presence of functionally significant stenosis using an SVM classifier based on the extracted features.

3- Experiments & Results:
- The study includes consecutively acquired CCTA scans of 166 patients who underwent invasive FFR measurements.
- Quantitative evaluation of LV myocardium resulted in an average Dice coefficient of 0.91 and an average mean absolute distance of 0.7 mm.
**Detection and Classification of Coronary Artery Plaque**

**1- Purposes:**
- Various types of atherosclerotic plaque and stenosis could lead to different coronary artery disease
- It is crucial to detect and classify the type of coronary artery plaque and determine the degree of coronary artery stenosis

**2- Methods:**
- First, a 3D convolutional neural network is utilized to extract features along the coronary artery.
- Subsequently, the extracted features are aggregated by a recurrent neural network that performs two simultaneous multiclass classification tasks.

**3- Experiments & Results**
- This study CCTA scans of 163 patients.
- For detection and characterization of coronary plaque, the method achieved an accuracy of 0.77
- And for detection stenosis it achieved an accuracy of 0.80

Overview of the proposed network [21]
Deep learning in Right Ventricle analysis[22]

1- Purposes:
• Currently RV segmentation is manually performed by clinical experts, which is lengthy, tiresome and sensitive to intra and interoperator variability
• This study aims to accurately segment the right ventricle (RV) from cardiac MRI using a fully automatic learning-based method

2- Methods
• CNN and stacked autoencoders are used for automatic detection and initial segmentation of the RV chamber
• First, in Step 1, the ROI containing the RV is determined in the image using a convolutional network trained to locate the RV.
• Then, in Step 2, the RV is initially segmented using a stacked-AE trained to delineate the RV.
• The obtained contour is used for initialization and incorporated into deformable models for segmentation in Step 3

3- Experiments & Results
• An average Dice metric of 82.5% along with an average Hausdorff distance of 7.85 mm were achieved for all the studied subjects

Endocardial contours of RV at ED from base to apex [22]

Block diagram of the integrated deep learning and deformable model algorithm [22]
**1- Purposes**

- Atrial fibrillation (AF) is a cardiac arrhythmia caused by abnormal electrical discharges in the atrium, often beginning with hemodynamic and/or structural changes in the left atrium (LA).
- Segmenting and measuring morphology of left atrium and proximal pulmonary veins from MR images.

**2- Methods**

- An encoder-decoder CNN architecture including 23 layers (11 in encoder, 12 in decoder units) is used.
- An adaptive fusion method is used to tackle the problem of non-isotropic spatial resolutions of a particular view.

**3- Experiments & Results**

- The method could achieve sensitivity (90%), specificity (99%), precision (94%), and efficiency levels of 10 seconds in GPU.

The overall view of proposed method [28]

After applying connected component, the biggest component is considered as trusted component and its volume is considered as a weight in fusion [28].
Coronary Arteries and Calcium Scoring with Deep Learning

- Coronary artery calcium segmentation
- Coronary Calcium Score Prediction
- Segmentation of Coronary Arteries from CT-scans
1- Purposes
- Coronary artery calcium (CAC) is a significant marker of atherosclerosis and cardiovascular events
- Present a system for the automatic quantification of calcium score in ECG-triggered non-contrast enhanced cardiac CT images

2- Methods
- CNN is used for the segmentation and classification of candidate lesions as coronary or not, previously extracted in the region of the heart using a cardiac atlas

3- Experiments & Results
- Lesions were detected with a sensitivity of 91%, a specificity of 95% and a PPV of 90%
- a) Correct coronary predictions. b) Correctly discarded lesion in the ascending aorta. c) Missed calcification. d) False positive on the ascending aorta.
1- Purposes
- The amount of calcium deposits in the coronary arteries is an important biomarker of cardiovascular disease.
- An automatic method based on fully-convolutional DNN is proposed to segment coronary calcium and predict Agatston score from any non-contrast chest CT.

3- Experiments & Results
Pearson correlation coefficient of 0.98 was achieved.

Examples of segmentation results:
(a) Accurate segmentation of calcifications
(b) Additional false-positive marked with a yellow arrow.

2- Methods
- The main idea underlying FCNNs for segmentation is extending a regular CNN in which a sequence of pooling operators progressively reduces the spatial size of the network, by adding successive layers where pooling operators are replaced by upsampling operators.

(a) Generic architecture of U-Net, (b) Generic architecture of a Fully Convolutional DenseNet [24]
1- Purposes

- Coronary artery disease diagnosis is today done by invasive methods, but research on using computational fluid dynamics to model the blood flow based on non-invasive imaging show great promise.
- A new method for refining the segmentation output to only include the coronary arteries is proposed.

2- Methods

- One network is trained on aorta segmentation and the other is trained on coronary artery segmentation.
- After obtaining a segmentation from the network trained on coronary arteries, the aorta segmentation is used to find the intersection between the aorta and the coronary arteries.
- This intersection is then used as start positions to select the coronary arteries and discarding spurious responses.

3- Experiments & Results

Dice Similarity of 60% was achieved by proposed method.
Multi substructure segmentation with Deep Learning

- Automatically Designing CNN
- Multi-Object Multi-Planar CNN (MO-MP-CNN)
**1- Purposes**

- A RL method is used to design an architecture of CNN automatically for medical image segmentation instead of trial-and-error method
- Hyper-parameters for the network can be optimized in an efficient manner
- The dataset consisting of 100 cine-MR images from ACDC challenge were used for training and evaluation

**2- Methods**

- A policy gradient reinforcement learning based method was used to learn the hyper-parameters of the baseline architecture (a densely connected encoder-decoder CNN)
- The hyperparameters of the network were considered as a policy which should be learned
- The learnable hyper-parameters were number of filters, filter height, filter width, and pooling layer.

**3- Experiments & Results**

The new architecture could beat the state-of-the-art methods with dice index of 0.88 and Hausdorff distance of 11 mm in average

![diagram](image.png)

optimally learned architecture[26]

Overview of proposed method[26]
**1- Purposes**

- Segmenting and measuring different substructures of cardiac from different modalities such as MRI and CT
- A dataset including 20 MR and 20 CT images from MMWHS challenge were used
- Substructures: myocardium, left atrium, left ventricle, right atrium, right ventricle, ascending aorta, and main pulmonary artery

**2- Methods**

- A multi-object multi-planar method based on the CNN were introduced
- Three different CNN were trained for each plane (sagittal, coronal, and axial)
- Then an adaptive fusion method were used to combine the output of each CNNs together
- The network consist of 24 layers and the Adam optimizer with learning rate of $10^{-4}$ was used to minimize cross entropy loss function

**3- Experiments & Results**

The precision and dice index of 0.93 and 0.90, and 0.87 and 0.85 were achieved for CT(50 seconds) and MR(17 seconds) images.

![Overview of proposed method][27]
## Challenges in Deep learning for Cardiac Image Analysis

<table>
<thead>
<tr>
<th>Challenges</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lack of enough data with different diseases and proper annotations</td>
</tr>
<tr>
<td>Difficulties in getting manual segmentation such as being tedious, being</td>
</tr>
<tr>
<td>prone to intra/inter observer error</td>
</tr>
<tr>
<td>Designing networks for analyzing cine-MR (4D) images directly, instead</td>
</tr>
<tr>
<td>of using 3D networks</td>
</tr>
<tr>
<td>Using information from different modalities (CT, PET, MRI, ...) in the</td>
</tr>
<tr>
<td>same network: Challenges in alignment/registration different modalities</td>
</tr>
<tr>
<td>Reduce the labor time for providing segmentation by choosing most</td>
</tr>
<tr>
<td>informative images/slices for annotation</td>
</tr>
<tr>
<td>Finding the optimum architecture for a given application</td>
</tr>
<tr>
<td>Data augmentation with learning algorithms</td>
</tr>
<tr>
<td>Motion correctness in imaging</td>
</tr>
<tr>
<td>Effectiveness of Deep Learning in Future of Cardiac Image Analysis</td>
</tr>
<tr>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td>Speeding up the diagnosis and analysis part by helping radiologists and physicians</td>
</tr>
<tr>
<td>Correction of Cardiac MR Motion from K-space [29]</td>
</tr>
<tr>
<td>Semi supervised learning to combine the power of the deep networks and radiologist in practice</td>
</tr>
<tr>
<td>Active deep learning can help effective and selective segmentation from slices of cine-cardiac MR images [30]</td>
</tr>
<tr>
<td>Using deep learning in cardiac image reconstruction from scanner</td>
</tr>
</tbody>
</table>
References


6) http://colah.github.io/posts/2015-08-Understanding-LSTMs/


10) https://skymind.ai/wiki/deep-reinforcement-learning


References


