ECG Hearbeat Classification Based on Multi-Scale Convolutional Neural Networks*

Ondrej Rozinek[1] [0000−0003−2248−7940] and Petr Dolezel[1] [0000−0002−7359−0764]

University of Pardubice, Studentska 95, 532 10 Pardubice, Czech Republic
ondrej.rozinek@gmail.com, petr.dolezel@upce.cz
https://fei.upce.cz/

Abstract. Clinical applications require automating ECG signal processing and classification. This paper investigates the impact of multiscale input filtering techniques and feature map blocks on the performance of CNN models for ECG classification. We conducted an ablation study using the AbnormalHeartbeat dataset, with 606 instances of ECG time series divided into five classes. We compared five multiscale input filtering techniques and four multiscale feature map blocks against a base model and non-multiscale input. Results showed that the combination of mean filter for multiscale input and residual connections for multiscale block achieved the highest accuracy of 64.47%. Residual connections were consistently effective across different filtering techniques, highlighting their potential to enhance CNN model performance for ECG classification. These findings can guide the design of future CNN models for ECG classification tasks, with further experimentation needed for optimal combinations in specific applications.

Keywords: ECG classification · deep learning · multiscale CNN · convolutional neural networks.

1 Introduction

CNNs (Convolutional Neural Networks) have been shown to be state of the art in many computer vision applications due to their ability to learn complex patterns in images with high accuracy [10, 12, 25]. CNNs achieve this by using convolutional layers that extract and combine local features from the input image, followed by fully connected layers that use the learned features to make a prediction [13]. Additionally, CNNs can be trained end-to-end using backpropagation, which allows them to automatically learn the optimal set of weights that minimize the prediction error [23]. This makes CNNs highly effective for a wide range of image recognition tasks, including object detection [21], image classification [25], and semantic segmentation [17].

* The work has been supported by the SGS grant no. SGS_2023_016 at the Faculty of Electrical Engineering and Informatics, University of Pardubice, Czech Republic. This support is very gratefully acknowledged.

CNNs have demonstrated significant improvements over traditional machine learning approaches in many computer vision tasks, especially in object recognition tasks such as the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [24]. The winning entries in the ILSVRC from 2012 to 2015 were all CNN-based models [10, 12, 25, 27]. Moreover, CNN-based models have set new records in various other image recognition tasks, including face recognition [28], image captioning [29], and visual question answering [2].

Overall, the effectiveness of CNNs can be attributed to their ability to learn and extract features from images in an end-to-end manner, which allows them to handle complex visual patterns and large datasets. With continued improvements in model architecture, optimization techniques, and hardware, CNNs are expected to continue to push the state of the art in computer vision applications.

2 Multiscale CNN and Related Work

Multiscale CNNs process signals at multiple resolutions, capturing fine-grained details and larger structural features. They can handle input data that varies in scale, leading to improved accuracy and reduced overfitting. Three types of multiscale CNNs are: (1) downsampled inputs, (2) multiple receptive field inputs, and (3) feature pyramid inputs [6, 8, 15, 16, 22, 26].

Inception network [26] uses convolutional layers with different filter sizes for processing images at multiple scales. U-Net [22] is designed for biomedical image segmentation tasks. Multiscale CNNs for time series classification [6] use convolutional layers at multiple scales to extract features. Pyramid Scene Parsing Network (PSPNet) [15] is for semantic segmentation of high-resolution images. Feature Pyramid Network (FPN) [16] is for object detection. Dual Attention Network (DANet) [8] uses self-attention mechanisms for image segmentation.

Relevant scientific competitions include ILSVRC, Kaggle, COCO Detection Challenge, VisDA, and Robust Vision Challenge.

3 CNN Classification of Electrocardiogram

Electrocardiograms (ECGs) are non-invasive tests that record the heart's electrical activity, with the PQRST complex being key for diagnosing cardiac conditions. Abnormalities in this complex can indicate arrhythmias, ischemia, or infarction. Research in ECG analysis focuses on developing algorithms to better detect and diagnose these abnormalities, often using machine learning models.

ECG classification is challenging due to signal variability, large dataset requirements, and interpretability. Convolutional neural networks (CNNs) dominate state-of-the-art methods, with transfer learning and data augmentation techniques improving performance. Publicly available datasets, like PhysioNet/CinC Challenge datasets [1], facilitate research. Recent studies propose hybrid deep learning models, such as CNN-RNNs [14] and convolutional attention models [18], achieving high accuracy in atrial fibrillation detection and PTB Diagnostic ECG Database tasks.

Further research is needed to address challenges in designing accurate, interpretable ECG classification models for clinical use.

4 Methodology

4.1 Multiscale Inputs

Our methodology utilizes multiscale input filtering techniques to preprocess a set of 1D electrocardiogram (ECG) signals. The primary goal of our methodology is to evaluate the impact of different multiscale filtering techniques on the performance of convolutional neural network (CNN) architectures for ECG classification. To ensure a fair comparison, all methods described in this section will have the same input and output formats, with the input being a preprocessed ECG signal and the output being the predicted class label. The input matrix $X \in \mathbb{R}^{N \times M}$ consists of separate signals in each row. We aim to develop a multiscale representation of these signals by filtering and downsampling them. The filtered and downsampled matrix $y(j) \in Y_{downsampled}$ is computed for each downsampling factor f_i using different filtering techniques.

In this work, we will compare five different multiscale input filtering techniques (subsampling, mean filter, Gaussian filter, bilateral filter and wavelet-based downsampling and denoising method).

Additionally, we will include the original (non-multiscale) 1D signal as a baseline for comparison. This "Original" method will not apply any filtering or downsampling to the input ECG signals, preserving the raw data for the CNN architecture. By including the original signal, we can evaluate the performance improvements, if any, that the multiscale filtering techniques provide over the raw input data.

Subsampling Subsampling is performed by selecting every f_i-th sample from the input signal x:

$$y(j) = x(j \cdot f_i), \tag{1}$$

where j is the index within the downsampled signal. This process is applied for various downsampling factors $f_1, f_2, \ldots, f_i, \ldots f_n$ to create a multiscale representation of the input signal.

Mean Filter The Mean Filter technique averages neighboring samples within the input signal $x \in \mathbb{R}^M$:

$$y(j) = \frac{1}{f_i} \sum_{n=0}^{f_i - 1} x(j \cdot f_i + n), \tag{2}$$

where j is the index within the downsampled signal. This process is repeated for different downsampling factors to obtain a multiscale representation of the input signal.

Gaussian Filter The Gaussian filter smooths the input signal x_i with a Gaussian function:

$$y(j) = \frac{1}{\sqrt{2\pi}\sigma} \sum_{n=-\infty}^{\infty} x_i(j+n) e^{-\frac{n^2}{2\sigma^2}}, \quad (3)$$

where σ is the standard deviation of the Gaussian filter. The filtered signal is then downsampled with a step size equal to the downsampling factor f_i. This process is executed for various downsampling factors to create a multiscale representation of the input signals.

Bilateral Filter The 1D Bilateral filter [19] is a non-linear, edge-preserving filter that smooths signals while preserving sharp edges. Given an input signal $x \in \mathbb{R}^M$, the output signal y is obtained by applying the 1D bilateral filter to each point j:

$$y(j) = \frac{1}{W} \sum_{n=-\lfloor \frac{k}{2} \rfloor}^{\lfloor \frac{k}{2} \rfloor} x(j+n) \cdot g_c(x(j+n) - x(j)) \cdot g_s(n), \quad (4)$$

where W is the normalization term, k is the kernel size, $g_c(\cdot)$ is the color Gaussian function with standard deviation σ_c, and $g_s(\cdot)$ is the spatial Gaussian function with standard deviation σ_s. The color Gaussian function and the spatial Gaussian function are defined as follows:

$$g_c(\Delta x) = e^{-\frac{(\Delta x)^2}{2\sigma_c^2}}, \quad (5)$$

$$g_s(n) = e^{-\frac{n^2}{2\sigma_s^2}}. \quad (6)$$

Wavelet-based Downsampling and Denoising Method The wavelet-based method for downsampling and denoising [7] time series data is built upon the discrete wavelet transform (DWT), thresholding techniques, and signal reconstruction.

Given a time series data X, the DWT decomposes the signal into a set of wavelet coefficients as follows:

$$X = A_n + D_n + D_{n-1} + \cdots + D_1, \quad (7)$$

where A_n is the approximation coefficients at level n, and D_i represents the detail coefficients at level i. The DWT is performed using a chosen wavelet function, such as the Daubechies wavelet (e.g., 'db4'), which offers a balance between smoothness and compact support.

The wavelet coefficients are thresholded to denoise the data. The threshold value (T) is computed as:

$$T = k \cdot \text{median}(|D_n|), \tag{8}$$

where k is the downsample factor, and $|D_n|$ is the absolute value of the detail coefficients at the highest level of decomposition. The threshold is applied using either a soft or hard thresholding mode. Soft thresholding is defined as:

$$Y_i = \text{sign}(D_i) \cdot \max(0, |D_i| - T), \tag{9}$$

where Y_i is the thresholded coefficient, and $\text{sign}(D_i)$ represents the sign of the detail coefficient D_i.

Hard thresholding is defined as:

$$Y_i = D_i \cdot I_{(|D_i|>T)}, \tag{10}$$

where $I_{(|D_i|>T)}$ is an indicator function that equals 1 if the condition is true and 0 otherwise. In our experiment we use only soft thresholding.

The denoised signal is reconstructed using the inverse discrete wavelet transform (IDWT) as follows:

$$X_{\text{denoised}} = \text{IDWT}(A_n, Y_n, Y_{n-1}, \ldots, Y_1), \tag{11}$$

where $x(j) \in X_{\text{denoised}}$ represents the denoised signal, and Y_i corresponds to the thresholded detail coefficients at level i.

Finally, the subsampling is performed by selecting every f_i-th sample from the input signal x:

$$y(j) = x(j \cdot f_i), \tag{12}$$

where j is the index within the downsampled signal. This process is applied for various downsampling factors $f_1, f_2, \ldots, f_i, \ldots f_n$ to create a multiscale representation of the input signal.

4.2 Multiscale Feature Maps

In this section, we explore the effects of incorporating different multiscale feature map blocks in our CNN models for ECG classification. We investigate four distinct types of multiscale blocks, each designed to capture different aspects of the input data. These blocks are compared against a base model that serves as a foundation for our ablation study.

Through the ablation study, we aim to quantify the contribution of the each block to the overall performance of our model by removing it and comparing the results with and without the block.

Base Model We present a base model that serves as a starting point for our ablation study. The base model consists of three 1-D convolutional layers with different filter sizes (3, 5, and 7) and a hyperbolic tangent (tanh) activation function. No pooling or residual connections are employed in this base model. The base model serves as a foundation for further investigation and comparison with other more complex models that incorporate techniques such as multiscale blocks, residual connections, and different types of convolutions.

Multiscale Block Inception Convolution In this ablation study, we investigate the impact of a used block known as the multiscale block inception convolution (further Inception Convolution) on the performance of our model. The Inception Convolution is inspired by the inception module introduced by [26] in their seminal work on the GoogLeNet architecture. The inception module aims to capture various spatial and channel-wise patterns within the input tensor by employing different filter sizes in parallel.

The Inception Convolution consists of three convolutional layers with different filter sizes (1, 3, and 5) and a max-pooling layer with a pool size of 3. The output feature maps of these layers are concatenated to form a combined feature map.

Multiscale Block with Depthwise Separable Convolution We investigate the impact of another used block known as the multiscale block with depthwise separable convolution [11] (further Depthwise Convolution) on the performance of our model. The Depthwise Convolution employs depthwise separable convolutions, which factorize a standard convolution into a depthwise convolution followed by a point-wise convolution, thus reducing the number of parameters and computational cost.

The Depthwise Convolution consists of three depthwise separable convolutional layers with different filter sizes (3, 5, and 7). The resulting feature maps from these layers are merged also together to create a unified feature map.

Multiscale Block with Dilated Convolution We examine the influence of an alternative building block, known as the multiscale block with dilated convolution [30] (Dilated Convolution), on our model's performance. The Dilated Convolution employs dilated convolutions, which introduce a dilation factor to increase the receptive field of the convolutional layers without increasing the number of parameters or computational cost.

The Dilated Convolution consists of three convolutional layers with different filter sizes (3, 5, and 7) and a dilation rate of 2. The output feature maps of these layers are concatenated to form a combined feature map.

Multiscale Block with Residual Connections We explore the effect of incorporating an additional block known as the multiscale block with residual connections [10] (also referred to as Residual Connections) on our model's performance.

The Residual Connections employs residual connections, which are a technique to mitigate the vanishing gradient problem in deep networks by allowing the gradients to flow through skip connections, thus improving the model's training and performance.

The Residual Connections consists of three pairs of convolutional layers with different filter sizes (3, 5, and 7) and a residual connection for each pair. The resulting feature maps from these layers are combined to create a unified feature map.

4.3 Training Details

We use the Adam optimizer for the training of the models, minimizing a categorical cross entropy loss function. The models are trained for a maximum of 10,000 epochs with early stopping, which monitors the validation loss and has a patience of 10 epochs. We use a validation split of 0.2, meaning that 20% of the training dataset is reserved for validation purposes.

In our methodology, we employ the Glorot Uniform initializer (also known as Xavier Uniform initializer) for weight initialization in our 1D convolutional layers. The Glorot Uniform initializer is designed to maintain a specific variance in the activations of the neurons, which helps avoid vanishing or exploding gradients during training.

For each model, we carry out 10 training sessions to mitigate the stochasticity of the training experiment. In each session, the dataset is randomly split into training and validation subsets, and the best model validated on the validation set is saved at the end of the training session. The performance of the trained models is then evaluated on the test set.

5 Experimental Results

In our ablation study, we used the AbnormalHeartbeat dataset for the task of classifying heartbeat recordings into one of five classes. This dataset contains a total of 606 instances, with each instance being a time series of length 3,053. The time series represent the change in amplitude over time during an examination of patients suffering from common arrhythmias. The dataset is divided into training and testing sets, each containing 303 instances.

The AbnormalHeartbeat dataset was obtained from a combination of sources, including the iStethoscope Pro iPhone app and clinical trials using the digital stethoscope DigiScope. All instances were resampled to 4,000Hz and truncated to the shortest instance length. The original data can be found at the provided link, and the original paper is by [3].

The dataset is composed of five classes:

– **Artifact** (40 cases): These are recordings that contain noise or other artifacts that can obscure the true nature of the heartbeat. They are not indicative of

any particular disease but are important to recognize in order to distinguish them from pathological cases [9].
- **ExtraStole** (46 cases): This class corresponds to recordings of heartbeats with premature ventricular contractions (PVCs) or extra systoles. PVCs are extra heartbeats that disrupt the normal rhythm of the heart and can be caused by various factors such as stress, caffeine, and heart diseases [31].
- **Murmur** (129 cases): Murmurs are abnormal heart sounds caused by turbulent blood flow across the heart valves. They can be indicative of various heart conditions, such as valve stenosis, regurgitation, or congenital heart defects [4].
- **Normal** (351 cases): This class consists of recordings of normal heartbeats, which exhibit a regular rhythm and no abnormal sounds. A normal heartbeat typically includes two main sounds: the first (S1) and second (S2) heart sounds, caused by the closure of the atrioventricular and semilunar valves, respectively [5].
- **ExtraHLS** (40 cases): This class corresponds to recordings with extra heart lung sounds, which are additional heart sounds that can be indicative of certain cardiac conditions, such as heart failure or pericarditis.

Algorithm 1 Ablation study of multiscale inputs and multiscale feature map blocks

```
DownFactors ← [8, 4, 2, 1]
Results ← []
Reps ← 10
for each InputFunc in InputFuncs do
    for each MultiFunc in MultiFuncs do
        for i in Reps do
            TrainData ← InputFunc(X_train, DownFactors)
            TestData ← InputFunc(X_test, DownFactors)
            Model ← main_model(MultiFunc)
            Compile Model with loss, optimizer, and metrics
            Define early stopping callback
            Fit Model with TrainData, y_train and validation split
            Score ← Model.evaluate(TestData, y_test)
        Compute accuracy and loss
        Store results in Results
```

As shown in Algorithm 1, we conducted an ablation study to compare the performance of different multiscale input filtering techniques and multiscale feature map blocks. The default train-test split was created through a random partition. In our ablation study, we investigated the impact of incorporating different multiscale feature map blocks in our CNN models for ECG classification. We analyzed four distinct types of multiscale blocks against a base model. By comparing the performance of each model, we aimed to determine the best multiscale input block and feature map block for ECG classification.

Table 1. Experimental results sorted by accuracy evaluated over testing set (descending)

Multiscale Inputs	Multiscale Block	Max Accuracy (%)
Mean Filter	Residual Connections	64.47
Gaussian Filter	Residual Connections	63.49
Wavelet Denoising	Residual Connections	63.49
Bilateral Filter	Residual Connections	63.16
Gaussian Filter	Base Model	63.16
Bilateral Filter	Base Model	62.17
Subsampling	Residual Connections	62.17
Subsampling	Base Model	61.84
Bilateral Filter	Inception Convolution	61.51
Gaussian Filter	Inception Convolution	61.18
Original	Base Model	61.18
Mean Filter	Base Model	60.86
Gaussian Filter	Depthwise Convolution	60.86
Mean Filter	Dilated Convolution	60.53
Gaussian Filter	Dilated Convolution	60.53
Subsampling	Depthwise Convolution	60.53
Mean Filter	Depthwise Convolution	60.20
Original	Residual Connections	59.87
Subsampling	Dilated Convolution	59.87
Bilateral Filter	Depthwise Convolution	59.54
Subsampling	Inception Convolution	59.54
Wavelet Denoising	Dilated Convolution	59.87
Wavelet Denoising	Base Model	60.86
Mean Filter	Inception Convolution	58.22
Bilateral Filter	Dilated Convolution	58.88
Original	Inception Convolution	57.89
Original	Depthwise Convolution	57.89
Original	Dilated Convolution	57.89
Wavelet Denoising	Inception Convolution	57.89
Wavelet Denoising	Depthwise Convolution	57.89

Please refer to the experimental results table for a detailed comparison of the performance of each model in terms of accuracy. In Table 1 we can identify the best multiscale input block and feature map block for this task.

6 Discussion and Conclusion

In this study, we investigated the impact of different multiscale input filtering techniques and multiscale feature map blocks on the performance of CNN architectures for ECG classification. We evaluated five multiscale input filtering techniques (subsampling, mean filter, Gaussian filter, bilateral filter, and wavelet-based downsampling and denoising method) and four multiscale feature map blocks (base model, depthwise convolution, inception convolution, and different filters) using the AbnormalHeartbeat dataset.

Our experimental results showed that the combination of the mean filter for multiscale input and the residual connections for multiscale block provided the best performance with an accuracy of 64.47%. The residual connections consistently achieved higher accuracy across different multiscale input filtering techniques, highlighting the effectiveness of using residual connections in the multiscale blocks for enhancing the performance of CNN models for ECG classification.

It is worth noting that the results are specific to the dataset and problem at hand. Therefore, further experimentation with different datasets and ECG classification tasks may yield different optimal combinations of multiscale input filtering techniques and multiscale feature map blocks. Additionally, future research can explore other filtering techniques and multiscale block designs to further improve the performance of CNN models for ECG classification. An experiment was also performed on this data in the paper [20]

In conclusion, our study demonstrated the importance of multiscale analysis in ECG classification and provided insights into the effectiveness of different multiscale input filtering techniques and multiscale feature map blocks in CNN architectures. Our findings can be useful for designing effective CNN models for ECG classification, which can have practical applications in the field of cardiology, such as in automated ECG diagnosis systems for early detection of arrhythmias and other cardiac conditions.

Acknowledgment

We would like to thank the software technology company Rozinet s.r.o. (www.rozinet.net) without whose support, this article would not have been possible to create by providing technical resources and space for research.

References

1. Physionet/cinc challenge 2017. `https://physionet.org/content/challenge-2017/`, accessed: 2023-03-08
2. Antol, S., Agrawal, A., Lu, J., Mitchell, M., Batra, D., Zitnick, C.L., Parikh, D.: Vqa: Visual question answering. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 2425–2433 (2015)
3. Bentley, P., Nordehn, G., Coimbra, M., Mannor, S., Getz, R.: The pascal classifying heart sounds challenge 2011 (chsc2011) results (2011)
4. Bonow, R.O., Otto, C.M.: Valvular Heart Disease: A Companion to Braunwald's Heart Disease. Elsevier (2020)
5. Boron, W.F., Boulpaep, E.L.: Medical physiology E-book. Elsevier Health Sciences (2016)
6. Cui, Z., Chen, W., Chen, Y.: Multi-scale convolutional neural networks for time series classification. arXiv preprint arXiv:1603.06995 (2016)

7. Donoho, D.L.: De-noising by soft-thresholding. IEEE Transactions on Information Theory **41**(3), 613–627 (1995)
8. Fu, J., Liu, J., Tian, H., Li, Y., Bao, Y., Fang, Z., Lu, H.: Dual attention network for scene segmentation. IEEE Transactions on Image Processing **29**, 7873–7887 (2020)
9. Grady, D.: Manual of Medical-Surgical Nursing Care: Nursing Interventions and Collaborative Management. Elsevier Health Sciences (2010)
10. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778 (2016)
11. Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H.: Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861 (2017)
12. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems. pp. 1097–1105 (2012)
13. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. Proceedings of the IEEE **86**(11), 2278–2324 (1998)
14. Li, X., Li, F., Zhou, X., et al.: A hybrid cnn-rnn model for atrial fibrillation detection from ecg recordings. In: Proceedings of the 2018 Computing in Cardiology Conference. pp. 1–4 (2018)
15. Li, Z., Peng, C., Yu, G., Zhang, X., Deng, Y.: Pyramid scene parsing network. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 2881–2890 (2017)
16. Lin, T.Y., Goyal, P., Girshick, R., He, K., Dollar, P.: Focal loss for dense object detection. In: Proceedings of the IEEE international conference on computer vision. pp. 2980–2988 (2017)
17. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 3431–3440 (2015)
18. Ozturk, S., Yildirim, O., Kocadag, I., et al.: A novel 1d convolutional attention model for ecg classification. Journal of Medical Systems **44**(7), 1–10 (2020)
19. Paris, S., et al.: Bilateral filtering: Theory and applications. Foundations and Trends in Computer Graphics and Vision **4**(1), 1–73 (2009)
20. Raza, A., Mehmood, A., Ullah, S., Ahmad, M., Choi, G.S., On, B.W.: Heartbeat sound signal classification using deep learning. Sensors **19**(21), 4819 (2019)
21. Ren, S., He, K., Girshick, R., Sun, J.: Faster r-cnn: Towards real-time object detection with region proposal networks. In: Advances in neural information processing systems. pp. 91–99 (2015)
22. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention. pp. 234–241. Springer (2015)
23. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning representations by back-propagating errors. Nature **323**(6088), 533–536 (1986)
24. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Fei-Fei, L.: Imagenet large scale visual recognition challenge. International Journal of Computer Vision **115**(3), 211–252 (2015)
25. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint (2014), arXiv:1409.1556

26. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1–9 (2014)
27. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z.: Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1–9 (2015)
28. Taigman, Y., Yang, M., Ranzato, M., Wolf, L.: Deepface: Closing the gap to human-level performance in face verification. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1701–1708 (2014)
29. Xu, K., Ba, J., Kiros, R., Cho, K., Courville, A., Salakhutdinov, R., Zemel, R., Bengio, Y.: Show, attend and tell: Neural image caption generation with visual attention. In: International Conference on Machine Learning. pp. 2048–2057 (2015)
30. Yu, F., Koltun, V.: Multi-scale context aggregation by dilated convolutions. arXiv preprint arXiv:1511.07122 (2015)
31. Zipes, D.P., Jalife, J.: Cardiac Electrophysiology: From Cell to Bedside. Elsevier Health Sciences (2018)