

Encoding Time Series as Multi-Scale Signed Recurrence Plots for Classification Using Fully Convolutional Networks

时间序列的分类任务近来已在深度学习方面取得了较大的进展。其中一种有较大发展空间的方式就是将时间序列编码为递归图，以便利用最新的DNN来提高时间序列分类的准确性。并且大量的研究也已经证明这种方法却得了令人瞩目的成果。然而，如何处理独特的区域规模（distinctive region scale）和时间序列的长度以及趋势混淆（tendency confusion）仍然是一个待解决的问题。

在本文，作者对递归图进行了改进，得到了一个新的递归图表达方式——多尺度递归图（MS-RP: Multi-scale Signed Recurrence Plots），并且还提出了一个基于MS-RP图像和完全卷积神经网络FCN对时间序列分类的新方法。这个方法首先引入相位空间维数 m 和RP的时延嵌入 t ，以生成多尺度的RP图像。然后，通过使用非对称结构，构造RP图像以表示很长的序列（包含700个以上的时间戳）。接下来，通过将设计的符号掩码相乘获得MS-RP图像，以消除趋势混淆。最后，用MS-RP图像训练FCN进行分类。

Literature Review

在以往的研究中，TSC方法主要分为三类：feature-based、distance-base和ensemble based。基于特征的方法是从时间序列中提取出具有代表性的特征，然后使用分类器将每一个时间序列映射到它对应的类别中；基于距离的方法是通过测量测试集与训练集之间的弹性距离，根据距离度量相似度，为时间序列分配标签；基于集合的方法将不同的特征和多个分类器集成在一个框架中，从而获得互补的效果和更好的分类精度。

为了更好地利用CNN的出色分类能力，最近的一些工作首先将时间序列编码为图像，然后将TSC问题转换为图像分类。之前提到的《*Classification of time-series images using deep convolutional neural networks*》就通过将时间序列编码为递归图，然后在RP图像上训练CNN以执行分类。这项工作在标准数据集上获得了最佳的分类性能，引起了学界的兴趣。

在本片论文中，收到RP提供的丰富伦理信息以及DNN在图像分类中杰出成果的启发，论文将递归图和TSC最新的DNN分类器整合到一个框架中。

Challenges & Solutions

论文在实践过程中总结了三个主要挑战，并提供了以下解决方案：

1、**挑战一：**不同数据集的序列的长度差异很大，并且它们独特的区域通常以不同的比例分布。

解决方案：现有方法只能通过调整图像大小来解决此问题。但是，为了避免高计算开销，将可调大小限制在较小范围内，这通常会降低RP的表示能力。我们通过额外引入RP的相空间维数（ m ）和嵌入时间延迟（ τ ）来解决这一挑战，这两个值在其他论文提到的方法中通常是固定的。而我们的论文根据不同的数据集探索不同的 m 、 τ 和图像的尺寸，以丰富RP图像的比例。然后，具有合适比例的图像将被选作分类器的输入。

2、**挑战二**：RP不擅长编码非常长的序列，尤其是当序列的长度大于700时。对于非常长的序列，其RP图像的大小太大，以至于它们被降采样到足够小，这导致图像信息的丢失。

解决方案：我们通过构造不对称的RP图像来应对这一挑战。具体地，首先将序列分为两部分，并且将每一部分编码为RP图像。然后，由于RP图像结构的对称性，每个RP图像的斜三角矩阵被提取并重新组合为一个图像。通过构造不对称矩阵，明显减少了由于下采样引起的信息丢失。

3、**挑战三**：RP混淆了时间序列的趋势。原因在于，用于计算相空间中状态距离的范数运算使RP图像中的所有像素值都保持正值。因此，RP无法区分序列的上升和下降趋势。

解决方案：我们通过引入符号规则来应对这一挑战。具体来说，将计算设计的带符号掩码，这些带符号掩码将使用正值和负值来指示序列的上升和下降趋势。

Contribution

我们将上述解决方案整合在一起，以提出多尺度符号递归图（MS-RP），然后将FCN和ResNet（这是TSC的最新分类器）应用于对MS-RP图像进行分类。与最新的时间序列分类算法相比，该算法的优势主要体现在两个方面：

- 1、首先，论文提出的MS-RP保留了RP在提供时间相关性和放大独特区域方面的优势。而且，与其他图像编码算法相比，MS-RP更好地适应了趋势，长度和比例尺序列的变化。其次，最新的深度学习分类器FCN和ResNet用于处理转换后的MS-RP图像，从而进一步提高了分类性能。
- 2、论文提出的方法在45个UCR时间序列分类数据集中实现了卓越的性能，并且分层提供了验证实验。
- 3、此外，利用t分布随机邻居嵌入（**t-SNE**），论文可视化了网络学习到的潜在表示的空间分布。它清楚地表明，MS-RP在提取特征时更好地利用了DNN的优势。

Methodology

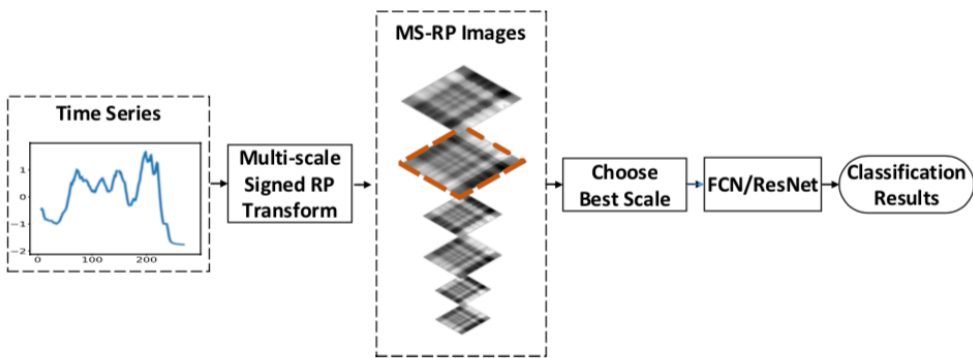


Figure 1. The framework of our approach. The x -axis and y -axis of time series represent the length and the amplitude, respectively.

论文提出的方法包括了两个阶段：

- 1、第一个阶段，论文将RP全面改进为MS-RP，并将时间序列编码为MS-RP图像。
- 2、第二个阶段，全连接卷积网络&残差网络处理这些时间序列编码的图像。

Proposed MS-RP

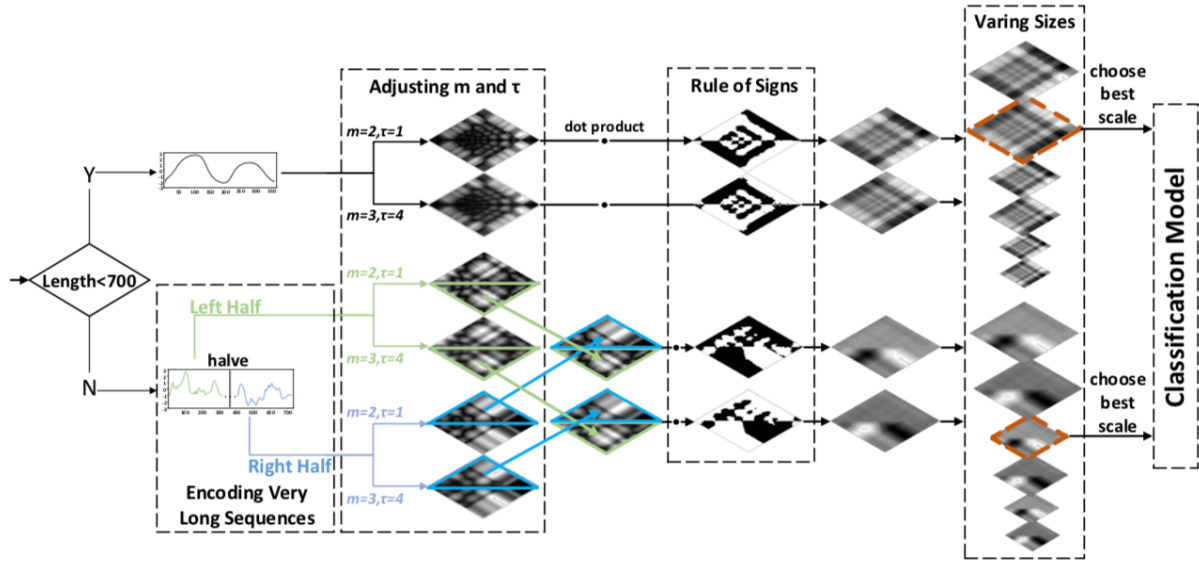


Figure 2. The architecture of our proposed MS-RP. The x -axis and y -axis of time series represent the length and the amplitude, respectively.

如图所示，我们将输入序列分为两种情况。对于短序列（长度小于700），首先将序列编码为两个不同比例的RP图像，然后提取符号掩码并将其乘以这些RP图像。然后，将这些多尺度带符号的RP图像调整为多种尺寸。最后，通过验证集上的分类性能，选择具有最佳比例的图片并将其作为输入，而不是原始序列进行分类。对于长序列（长度超过700个），首先将一个序列分成两个等长的片段，每个片段像一个短序列一样被编码为RP图像。然后，利用RP的对称结构，将对应于两个分割块的两个RP图像重新组合为一个非对称图像。编码过程的其余部分与短序列保持一致。

Multi-Scale RP: An Improvement of RP

时间序列的具有区分性的区域通常以各种比例出现，并且序列之间的长度差异很大。现有方法通过调整图像大小来适应这些变化。然而，考虑到计算成本，将这些图像尺寸控制在相对较小的范围内，从而限制了RP的表示能力。

RP图像的生成过程类似于空洞卷积（dilated convolution）运算的过程。滑动到原始数据上的子序列可以被视为空洞卷积核，只是范数计算被内部乘积代替。子序列的长度和采样间隔分别对应于卷积核大小和空洞率，并且可以通过 m 和 τ 进行控制。 m 和 τ 的不同值会改变滑动子序列的感受野（receptive fields），并且可以以各种比例构建时间相关性。因此，论文引入相位空间尺寸 m 和RP的嵌入时间延迟 τ 来解决这一挑战，在其他文章中的做法是始终忽略并保持它们的值固定不变。根据不同的数据集，将 m 和 τ 的值与图像大小一起调整以生成多尺度图像。通过适当地选择该多尺度图像，时间相关性可以建在合适的尺度，和图像尺寸能较好地适应序列的长度变化，以及该网络的感受野。（ m ， τ ）最常用的值是（2，1）或（3，4）。本文都采用了这两组值，分别对应于两种不同比例的RP图像。（ m ， τ ）的搜索范围如此之小，是由于我们最初的动机是验证调整这两个参数的重要性，而不是寻找最佳值。

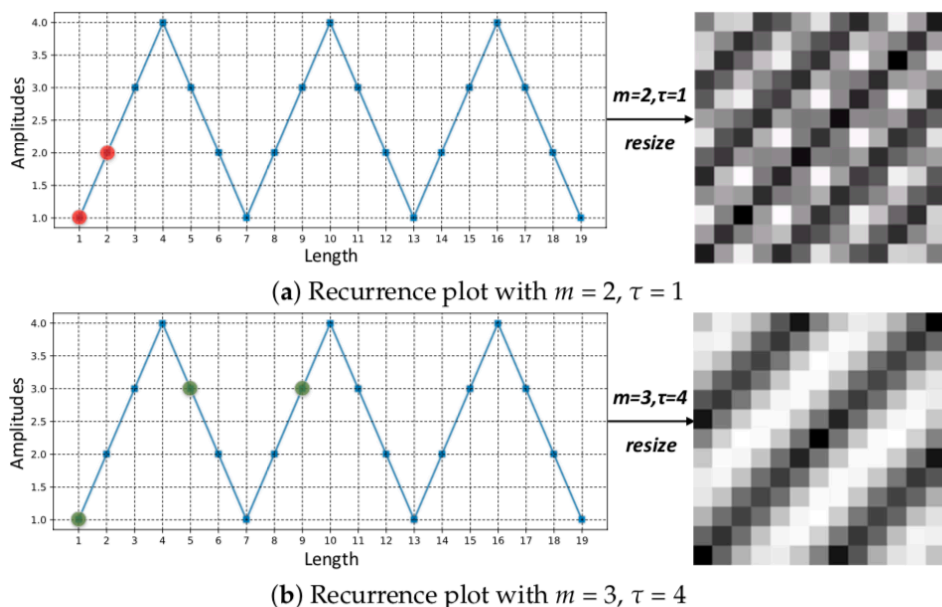


Figure 4. The subsequences and RP images of a triangular periodic sequence, with different values of m and τ . The red and blue dots in (a) and (b) correspond to the first subsequences of the sequence when (m, τ) are $(2, 1)$ and $(3, 4)$, respectively.

图4显示了具有这两组 m 和 τ 的三角形的周期序列及其RP图像。可以看出，即使具有相同的图像尺寸，较小的 (m, τ) 值也可以生成更精细的图像，而较大的 (m, τ) 值可以生成具有整体信息的图像。

Asymmetrical RP for Encoding Very Long Sequences

对于非常长的序列 (> 700)，RP图像的大小可能非常大。一方面，具有如此大尺寸的RP图像将带来计算爆炸性的增长；另一方面，如果将这些图像调整为合理大小，将导致严重的信息丢失。

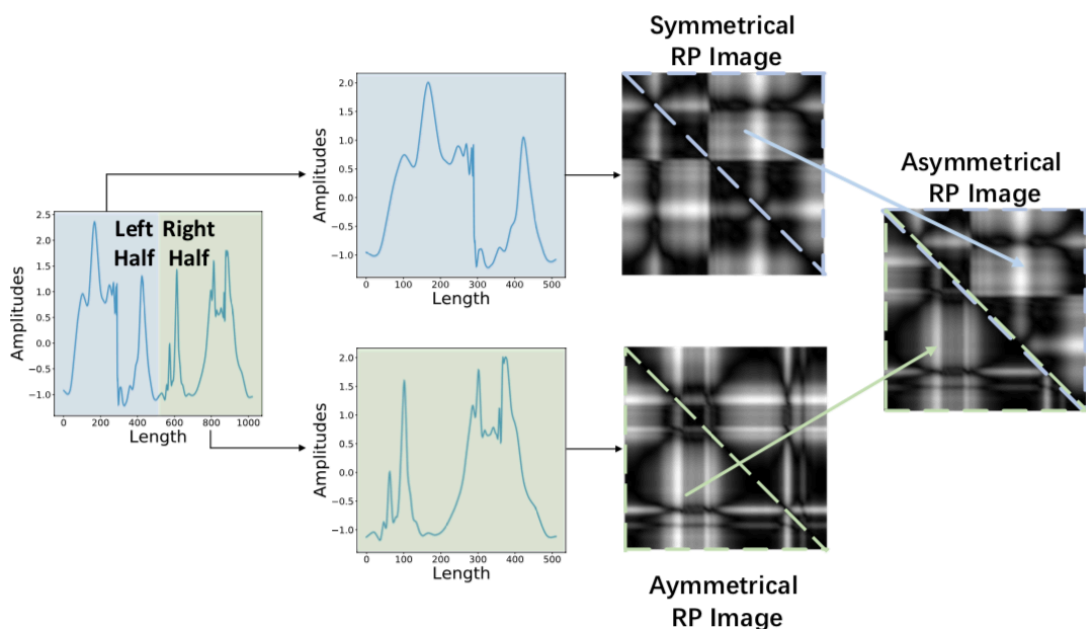


Figure 5. The illustration of constructing an asymmetrical RP image (the sequence comes from 'Mallat' dataset).

为了应对这一挑战，提出了非对称RP。图5显示了构造不对称RP图像的过程。如图所示，一个序列被分成两半，每一个被编码为图像。然后，利用RP的对称结构，分别提取两幅图像的上和下斜角，然后重新组合为一张图像。通过构造非对称RP图像，减轻了调整大小过程中带来的信息丢失，克服了对称RP的信息冗余问题。

Rule of Signs

$$RP_{i,j}(\epsilon) = \frac{\|\vec{x}(i) - \vec{x}(j)\|}{\|\vec{x}(i) - \vec{x}(j)\|}, \vec{x}(\cdot) \in \mathfrak{R}^m, i, j = 1, \dots, N \quad (2)$$

如等式 (2) 所示，将范数运算用于相空间中的状态之间的距离计算，这些距离对应于RP图像中的像素值。常用的范数运算为L1-范数，L2-范数和L ∞ -范数；然而，无论选择哪种规范运算，RP图像的所有像素值都是正值，从而导致严重的RP倾向混淆 (tendency confusion) 问题。

用一个简单的例子进行说明， s_1 和 s_2 表示两段短序列，它们的值分别为【1, 2, 3, 】和【3, 2, 1】，分别对应的是两个相反的单调趋势。等式 (2) 应用于RP矩阵的计算，其中 $\|\cdot\|$ 是L2范式，(m,t)的值为 (2, 1)，通过计算得到 s_1 和 s_2 的相对位置矩阵RP matrixes，如下所示：

$$RP_{s_1} = \begin{pmatrix} 0 & \sqrt{2} \\ \sqrt{2} & 0 \end{pmatrix}, \quad (3)$$

$$RP_{s_2} = \begin{pmatrix} 0 & \sqrt{2} \\ \sqrt{2} & 0 \end{pmatrix}. \quad (4)$$

可以看到， s_1 和 s_2 的相似位置矩阵都是一样的。

为了克服这个挑战，论文引入了相似位置矩阵RP的符号规则。首先，序列被映射到相空间中，然后进行相空间中状态之间的减法和范数运算，分别获得状态差矢量和RP图像像素值。其次，我们将每个状态差向量分别求和。然后，提取总和值的符号以构造具有与RP图像相同尺寸的符号掩模。最后，将符号掩码乘以RP图像，从而获得带符号的RP图像。整个过程由公式 (5) 定义：

$$RP_{i,j}(\epsilon) = \frac{\text{sum}(\vec{x}(i) - \vec{x}(j)) \cdot \|\vec{x}(i) - \vec{x}(j)\|}{|\text{sum}(\vec{x}(i) - \vec{x}(j))|}, \quad (5)$$

$$\vec{x}(\cdot) \in \mathfrak{R}^m, i, j = 1, \dots, N$$

如图6所示，图6（左）显示了具有相反趋势的两个序列的RP图像。这些序列来自“Synthetic Control”数据集。图6（中）显示了这两个序列的RP图像，很难区分它们。图6（右）显示了两个序列的带符号的RP图像，这些带符号的图像反映了序列的趋势，可以轻松区分。

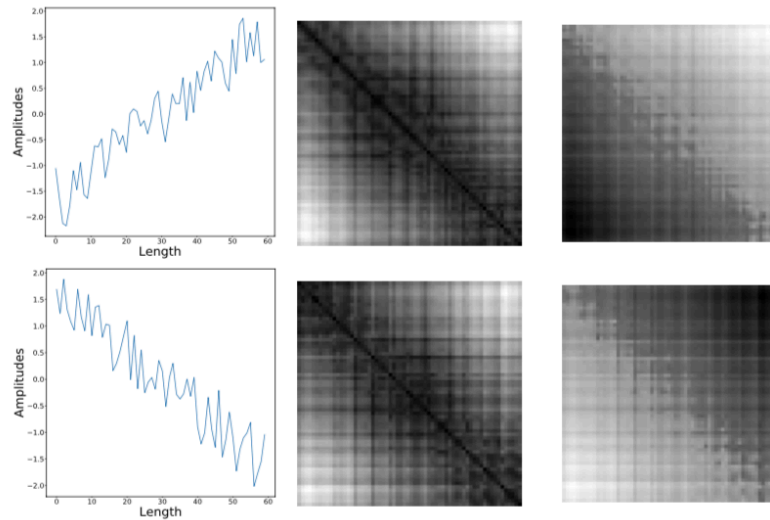


Figure 6. The RP images of sequences with opposite tendencies. (left column: two sequences with opposite tendencies from ‘SyntheticControl’ dataset, middle column: the RP images of the sequences, right column: signed RP images of the sequences).

Classification Using FCN on MS-RP Images

在最后一节中，RP被全面修改为MS-RP，以将时间序列编码为图像。高性能分类器应用于这些图像。现有方法通常将RP与k-最近邻（kNN）分类器或传统CNN分类器结合在一起。但是，kNN分类器的性能在很大程度上取决于手动设计的特征。另外，尽管传统的CNN分类器将特征学习和分类统一在一个框架中，但是合并操作会导致严重的信息丢失，并且具有大量参数的完全连接的层可能会导致过拟合的发生。

为了解决这些问题，本文引入了FCN和ResNet来处理MS-RP图像，根据图像数据格式将其扩展为2D版本。FCN和ResNet在中《*Multilevel wavelet decomposition network for interpretable time series analysis.*》首先被提出作为基线分类器，它们被广泛认为是TSC的最新分类器。这两个网络的体系结构如图7b，c所示。

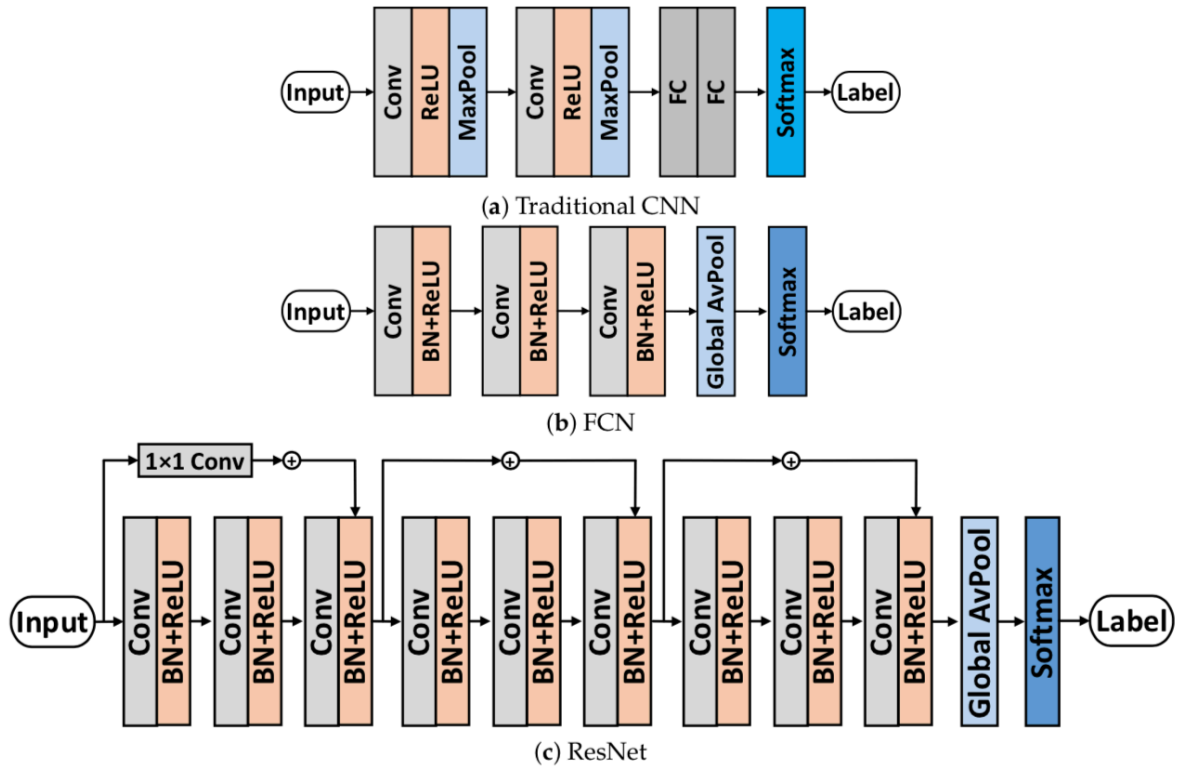


Figure 7. The architectures of network classifiers used in this paper.

FCN是一个完全卷积的网络，具有三个卷积层。每层都遵循一个批标准化（BN）层和一个ReLU激活函数。FCN没有完全连接（FC）层。卷积过程完成后，要素将通过全局平均池化层和Softmax层进行分类。ResNet通过残留连接扩展了FCN。它具有三个残差块，每个块的结构与FCN相同。

ResNet探索了具有更深架构的网络，这是平衡更好的表示形式和过度拟合的折衷方案。

Experiments and Analysis

论文在UCR档案的45个数据集上评估了我们提出的方法和最先进的竞争对手，该数据集是来自现实世界中各个领域的TSC数据集的集合。

不同算法之间的对比

论文进行的对比的方法有：

1、FCN和ResNet：这两个模型在《*Time series classification from scratch with deep neural networks: A strong baseline*》中提出，被认为是TSC的强基准和最佳基于DNN的分类器。

2、RP-CNN：《*Classification of time-series images using deep convolutional neural networks*》将RP与传统的CNN结合在一起，这与我们提出的方法类似。我们将其作为将时间序列编码为图像的方法的基准。

3、RP-FCN：此模型将RP和FCN合并为一个框架。它是我们提供的，用于RP和MS-RP之间的比较。

4、HIVE-COTE：该模型具有五种不同特征，具有各种异构分类器[《*Time series classification with HIVE-COTE: The hierarchical vote collective of transformation-based ensembles.*》]；在传统的时间序列分类方法中，它可以实现最先进的性能。

5、FCN残留分类流（FCN-RCF）：《*Multilevel wavelet decomposition network for interpretable time series analysis.*》通过细粒度小波将序列分解为子序列的多个频率，然后应用FCN处理这些子序列。该模型具有非常强的性能。

6、ALSTM-FCN：此模型在一个

框架中结合了FCN和ALSTM 《LSTM fully convolutional networks for time series classification.》。ALSTM补充了FCN的重要时间信息，从而明显提高了分类性能。所提出的ALSTM-FCN在基于DNN的方法中实现了最先进的性能。

MS-RP超参数的对比

在 (2, 1) 和 (3, 4) 之间选择 (m, τ) 的值，并且图像尺寸范围 (16、48、64、80、96、112、128)。

Result

Comparison between Different Values of m and τ

Table 3. Comparison in terms of error rates between different m and τ .

Dataset	Adiac	Face- All	Medical- Img	Mote- Strain	OSU- Leaf	CricketY	CricketZ	Fifty- Words	Lightning2	Lightning7
$m = 2, \tau = 1$	0.1379	0.0320	0.1947	0.0831	0.0165	0.2077	0.1897	0.2066	0.1475	0.1507
$m = 3, \tau = 4$	0.1637	0.0698	0.2316	0.1222	0.0620	0.1718	0.1615	0.1780	0.0984	0.1233

表3提供了十对分类错误率，用于在m和 τ 的两个不同组之间进行性能比较，图像尺寸保持一致。可以在表中找到两组错误率之间的明显差距。如上所述，(m, τ) 的不同值丰富了RP图像的比例，这有助于更好地表示时间序列。

Comparison between Symmetric RP and Asymmetric RP.

Table 4. Comparison in terms of error rates between symmetric RP and asymmetric RP.

Dataset	CinTorso	InlineSkate	Mallat	NonInThorax1	NonInThorax2	StarLigCurves
Asymmetric RP	0.1123	0.5491	0.0422	0.0361	0.0366	0.0180
Symmetric RP	0.2866	0.5382	0.0729	0.0539	0.0514	0.0232

论文建议使用非对称RP图像来编码很长的序列。为了比较对称和非对称RP的性能，选择了六个具有很长序列的UCR数据集，表4中提供了错误率。可以看出，非对称结构是有帮助的，尽管在CinCEGTorso的数据集中帮助很小。这可能是由于非对称RP更能够保留详细信息，而大多数选定的数据集都具有全局形状。

Comparison between Signed RP and Unsigned RP.

Table 5. Comparison in terms of error rates between signed RP and unsigned RP.

Dataset	CricketX	CricketY	CricketZ	Lightning7	OSU- Leaf	Syn- Control	Two- Patterns	UWaveX	UWaveY	UWaveZ
Signed RP	0.1821	0.1718	0.1615	0.1233	0.0165	0	0	0.1963	0.2725	0.2462
Unsigned RP	0.2187	0.2349	0.2064	0.1469	0.0744	0.2967	0.4850	0.3778	0.3425	0.3431

引入符号规则来克服RP的趋势混淆问题。为了评估其有效性，我们选择了10个数据集，并在表5中提供了带签名的和未签名的RP的性能。如图所示，带签名的RP获得了巨大的优势。因此，符号掩码是RP的有效补充，它描述了序列的趋势变化。实际上，符号掩码对于动作识别数据集更为关键，并且这些数据集的序列对趋势变化更为敏感。

Conclusion & Future Work

本文将RP全面改进为MS-RP，然后将TSC问题转化为DNN的图像分类任务。首先，引入相位空间维数 m 和RP的嵌入时延 τ 来丰富RP图像的尺度。其次，构造非对称RP来编码很长的序列。最后，引入符号法则来克服RP的趋势混淆问题。此外，对FCN和ResNet进行了训练以处理MS-RP图像，这是TSC的最新分类器。在45个UCR数据集上的实验结果表明，我们提出的方法优于最新技术，并且还通过验证实验对MS-RP的每个模块进行了分层验证。由于可穿戴式传感器的普及，我们的工作可以轻松地**扩展到实际应用**，例如运动识别，ECG健康以及手机上的睡眠状态监控。我们希望将这些有趣的工作作为未来的工作。