Aftershock Detection with Multi-Scale Description based Neural Network

Qi Zhang\textsuperscript{1,2}, Tong Xu\textsuperscript{1,*}, Hengshu Zhu\textsuperscript{2,*}, Lifu Zhang\textsuperscript{1}, Hui Xiong\textsuperscript{1,2,3}, Enhong Chen\textsuperscript{1}, Qi Liu\textsuperscript{1}

\textsuperscript{1}Anhui Province Key Lab of Big Data Analysis and Application, University of Science and Technology of China
\textsuperscript{2}Baidu Talent Intelligence Center, Baidu Inc,
\textsuperscript{3}Business Intelligence Lab, Baidu Research
Earthquake Events

Earthquake is one kind of worst nature disasters which may cause injury and loss of life and collapse of buildings.
Background

- **Earthquake Events**

  Earthquake is one kind of worst nature disasters which may cause injury and loss of life and collapse of buildings.

Earthquake events distribution all over the world in 2018

Nature Disaster: Earthquake
Background

- **Earthquake Events**

Earthquake is one kind of worst nature disasters which may cause injury and loss of life and collapse of buildings.
Background

- **Aftershock**
  Aftershocks refer to the smaller earthquakes that occur following large earthquakes, in the same area of the main shock.

  - Difficult to detect.
  - Collapse buildings that are damaged from the main shock.

![2016 Central Italy earthquake magnitude graph](image-url)
Aftershocks refer to the smaller earthquakes that occur following large earthquakes, in the same area of the main shock.

- Difficult to detect.
- Collapse buildings that are damaged from the main shock.
Background

- **Aftershock**
  Aftershocks refer to the smaller earthquakes that occur following large earthquakes, in the same area of the main shock.

- **Difficult to detect.**

- **Collapse buildings that are damaged from the main shock.**
Background

◆ **Aftershock**

Aftershocks refer to the smaller earthquakes that occur following large earthquakes, in the same area of the main shock.

◆ Difficult to detect.

◆ **Collapse buildings**
  that are damaged from the main shock.
Background

◆ Motivation

Benefit
◆ Automatic aftershock detection can support emergence actions.
◆ Useful for the research of geological activity and seismic mechanism.

Challenges
◆ Highly-noise and weak signal. How to effectively detect aftershocks?
   Multi-Scale Description based Neural Network.
◆ Multiple monitoring stations. How to utilize this relationship?
   multi-task learning strategy.
Background

◆ Motivation

Benefit
- Automatic aftershock detection can support emergence actions.
- Useful for the research of geological activity and seismic mechanism.

Challenges
- Highly-noise and weak signal. How to effectively detect aftershocks?
  *Multi-Scale Description based Neural Network.*
- Multiple monitoring stations. How to utilize this relationship?
  *multi-task learning strategy.*
Background

◆ Motivation

Benefit
◆ Automatic aftershock detection can support emergence actions.
◆ Useful for the research of geological activity and seismic mechanism.

Challenges
◆ Highly-noise and weak signal. How to effectively detect aftershocks?
  *Multi-Scale Description based Neural Network.*
◆ Multiple monitoring stations. How to utilize this relationship?
  *multi-task learning strategy.*
Background

◆ Motivation

Benefit

◆ Automatic aftershock detection can support emergence actions.

◆ Useful for the research of geological activity and seismic mechanism.

Challenges

◆ Highly-noise and weak signal. How to effectively detect aftershocks?
   *Multi-Scale Description based Neural Network.*

◆ Multiple monitoring stations. How to utilize this relationship?
   *multi-task learning strategy*.
Problem Formulation

- **Aftershock Detection**

**Definition:** Machine-learning-based Aftershock Detection.

Given a set of waveform windows $D$, where each $d_i \in D$ has a label $l_i$ for indicating the existence of seismic P-wave, the **objective** is to learn a predictive model $M$ for classifying waveform windows with respect to the label $y_i$. 

![Diagram showing waveform windows and classification process](Image)
Problem Formulation

- **Aftershock Detection**

Definition: Machine-learning-based Aftershock Detection.

Given a set of waveform windows $D$, where each $d_i \in D$ has a label $l_i$ for indicating the existence of seismic P-wave, the **objective** is to learn a predictive model $M$ for classifying waveform windows with respect to the label $y_i$. 

![Diagram](image)
**Problem Formulation**

- **Aftershock Detection**

  **Definition:** Machine-learning-based Aftershock Detection.

  Given a set of waveform windows $D$, where each $d_i \in D$ has a label $l_i$ for indicating the existence of seismic P-wave, the **objective** is to learn a predictive model $M$ for classifying waveform windows with respect to the label $y_i$. 
<table>
<thead>
<tr>
<th></th>
<th>Background</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Problem Formulation</td>
</tr>
<tr>
<td>3</td>
<td>Methodology</td>
</tr>
<tr>
<td>4</td>
<td>Experiments</td>
</tr>
<tr>
<td>5</td>
<td>Conclusion</td>
</tr>
</tbody>
</table>
Methodology

Data Set Description

- The waveforms after the Wenchuan M8.0 Earthquake.
- There are 2,833 aftershocks, corresponding to 9,891 pieces of seismic waveforms in short time window.
- These waveforms were recorded in three spatial dimensions (i.e., Z for the vertical channel, N for the north-south channel, and E for the east-west channel) by 15 monitoring stations.

![Example for aftershock detection.](image)

<table>
<thead>
<tr>
<th>Station</th>
<th>Number</th>
<th>Homology Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>JMG</td>
<td>1,208</td>
<td>1,208</td>
<td>100%</td>
</tr>
<tr>
<td>YZP</td>
<td>1,072</td>
<td>1,015</td>
<td>94.7%</td>
</tr>
<tr>
<td>QCH</td>
<td>894</td>
<td>890</td>
<td>99.6%</td>
</tr>
<tr>
<td>PUW</td>
<td>1,350</td>
<td>1,338</td>
<td>99.1%</td>
</tr>
<tr>
<td>WXZ</td>
<td>839</td>
<td>838</td>
<td>99.9%</td>
</tr>
<tr>
<td>SPA</td>
<td>574</td>
<td>574</td>
<td>100%</td>
</tr>
<tr>
<td>XJI</td>
<td>614</td>
<td>612</td>
<td>99.7%</td>
</tr>
<tr>
<td>HSH</td>
<td>821</td>
<td>821</td>
<td>100%</td>
</tr>
<tr>
<td>YGD</td>
<td>166</td>
<td>166</td>
<td>100%</td>
</tr>
<tr>
<td>JJS</td>
<td>908</td>
<td>903</td>
<td>99.4%</td>
</tr>
<tr>
<td>MXI</td>
<td>1,215</td>
<td>1,196</td>
<td>98.4%</td>
</tr>
<tr>
<td>XCO</td>
<td>223</td>
<td>223</td>
<td>100%</td>
</tr>
<tr>
<td>WDT</td>
<td>6</td>
<td>6</td>
<td>100%</td>
</tr>
<tr>
<td>MIA</td>
<td>1</td>
<td>1</td>
<td>100%</td>
</tr>
<tr>
<td><strong>SUM</strong></td>
<td><strong>9,891</strong></td>
<td><strong>9,791</strong></td>
<td><strong>99.0%</strong></td>
</tr>
</tbody>
</table>
Methodology

- **Data Set Description**
  - The waveforms after the **Wenchuan M8.0 Earthquake**.
  - There are 2,833 aftershocks, corresponding to 9,891 pieces of seismic waveforms in short time window.
  - These waveforms were recorded in three spatial dimensions (i.e., Z for the vertical channel, N for the north-south channel, and E for the east-west channel) by 15 monitoring stations.
Methodology

- **Data Set Description**
  - The waveforms after the Wenchuan M8.0 Earthquake.
  - There are 2,833 aftershocks, corresponding to 9,891 pieces of seismic waveforms in short time window.
  - These waveforms were recorded in three spatial dimensions (i.e., Z for the vertical channel, N for the north-south channel, and E for the east-west channel) by 15 monitoring stations.

<table>
<thead>
<tr>
<th>Station</th>
<th>Number</th>
<th>Homology Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>JMG</td>
<td>1,208</td>
<td>1,208</td>
<td>100%</td>
</tr>
<tr>
<td>YZP</td>
<td>1,072</td>
<td>1,015</td>
<td>94.7%</td>
</tr>
<tr>
<td>QCH</td>
<td>894</td>
<td>890</td>
<td>99.6%</td>
</tr>
<tr>
<td>PUW</td>
<td>1,350</td>
<td>1,338</td>
<td>99.1%</td>
</tr>
<tr>
<td>WXZ</td>
<td>839</td>
<td>838</td>
<td>99.9%</td>
</tr>
<tr>
<td>SPA</td>
<td>574</td>
<td>574</td>
<td>100%</td>
</tr>
<tr>
<td>XJI</td>
<td>614</td>
<td>612</td>
<td>99.7%</td>
</tr>
<tr>
<td>HSH</td>
<td>821</td>
<td>821</td>
<td>100%</td>
</tr>
<tr>
<td>YGD</td>
<td>166</td>
<td>166</td>
<td>100%</td>
</tr>
<tr>
<td>JJS</td>
<td>908</td>
<td>903</td>
<td>99.4%</td>
</tr>
<tr>
<td>MXY</td>
<td>1,215</td>
<td>1,196</td>
<td>98.4%</td>
</tr>
<tr>
<td>XCO</td>
<td>223</td>
<td>223</td>
<td>100%</td>
</tr>
<tr>
<td>WDT</td>
<td>6</td>
<td>6</td>
<td>100%</td>
</tr>
<tr>
<td>MIAX</td>
<td>1</td>
<td>1</td>
<td>100%</td>
</tr>
</tbody>
</table>

SUM 9,891 9,791 99.0%
Methodology

◆ Data Set Description

◆ The waveforms after the Wenchuan M8.0 Earthquake.

◆ There are 2,833 aftershocks, corresponding to 9,891 pieces of seismic waveforms in short time window.

◆ These waveforms were recorded in three spatial dimensions (i.e., Z for the vertical channel, N for the north-south channel, and E for the east-west channel) by 15 monitoring stations.

<table>
<thead>
<tr>
<th>Station</th>
<th>Number</th>
<th>Homology Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>JMG</td>
<td>1,208</td>
<td>1,208</td>
<td>100%</td>
</tr>
<tr>
<td>YZP</td>
<td>1,072</td>
<td>1,015</td>
<td>94.7%</td>
</tr>
<tr>
<td>QCH</td>
<td>894</td>
<td>890</td>
<td>99.6%</td>
</tr>
<tr>
<td>PUW</td>
<td>1,350</td>
<td>1,338</td>
<td>99.1%</td>
</tr>
<tr>
<td>WXZ</td>
<td>839</td>
<td>838</td>
<td>99.9%</td>
</tr>
<tr>
<td>SPA</td>
<td>574</td>
<td>574</td>
<td>100%</td>
</tr>
<tr>
<td>XJII</td>
<td>614</td>
<td>612</td>
<td>99.7%</td>
</tr>
<tr>
<td>HSH</td>
<td>821</td>
<td>821</td>
<td>100%</td>
</tr>
<tr>
<td>YGD</td>
<td>166</td>
<td>166</td>
<td>100%</td>
</tr>
<tr>
<td>JJS</td>
<td>908</td>
<td>903</td>
<td>99.4%</td>
</tr>
<tr>
<td>MXII</td>
<td>1,215</td>
<td>1,196</td>
<td>98.4%</td>
</tr>
<tr>
<td>XCO</td>
<td>223</td>
<td>223</td>
<td>100%</td>
</tr>
<tr>
<td>WDT</td>
<td>6</td>
<td>6</td>
<td>100%</td>
</tr>
<tr>
<td>MIAX</td>
<td>1</td>
<td>1</td>
<td>100%</td>
</tr>
</tbody>
</table>

SUM: 9,891 9,791 99.0%

(a) Example for aftershock detection.
Methodology

- **Characteristics of the Data**

  - STA/LTA (Short-Term Average/Long-Term Average) is the most widely-used earthquake detection approach.

  - The different scale-aware descriptions reflect different characteristics of seismic waveform.

Fig. 2: Some motivating examples of our multi-scale description based aftershock detection approach.
Methodology

- Characteristics of the Data
  - STA/LTA (Short-Term Average/Long-Term Average) is the most widely-used earthquake detection approach.
  - The different scale-aware descriptions reflect different characteristics of seismic waveform.

Fig. 2: Some motivating examples of our multi-scale description based aftershock detection approach.
Methodology

◆ Characteristics of the Data

We design a **Multi-Scale Description Based Neural Network** to extract appropriate features for improving the performance of aftershock detection.
MSD-cell: Generating Multi-Scale Description

The module needs to implement two key functions:

First function: extract multi-scale description
- The first function can remember prior features on different scales and add new scale feature.
- Implemented by a \(1\times3\times32/1\) convolutions

Second function: mix multi-scale description
- The second function can compare and mix these two kinds of features.
- Implemented by a \(1\times1\times32/1\) convolutions

Fig. 3: The detailed structure of MSD-cell, which can be expanded easily.
**MSD-cell: Generating Multi-Scale Description**

The module needs to implement two key functions:

1. **First function:** extract multi-scale description
   - The first function can remember prior features on different scales and add new scale feature.
   - Implemented by a $1 \times 3 \times 32/1$ convolutions

2. **Second function:** mix multi-scale description
   - The second function can compare and mix these two kinds of features.
   - Implemented by a $1 \times 1 \times 32/1$ convolutions

Fig. 3: The detailed structure of MSD-cell, which can be expanded easily.
Methodology

- **MSD-cell: Generating Multi-Scale Description**

  the module needs to implement two key functions

  **First function: extract multi-scale description**
  - The first function can remember prior features on different scales and add new scale feature.
  - Implemented by a $1 \times 3 \times 32 / 1$ convolutions

  **Second function: mix multi-scale description**
  - The second function can compare and mix these two kinds of features.
  - Implemented by a $1 \times 1 \times 32 / 1$ convolutions

Fig. 3: The detailed structure of MSD-cell, which can be expanded easily.
Methodology

- MSD-cell: Generating Multi-Scale Description
  - The module needs to implement two key functions
  - **First function: extract multi-scale description**
    - The first function can remember prior features on different scales and add new scale feature.
    - Implemented by a $1\times3\times32/1$ convolutions
  - **Second function: mix multi-scale description**
    - The second function can compare and mix these two kinds of features.
    - Implemented by a $1\times1\times32/1$ convolutions

Fig. 3: The detailed structure of MSD-cell, which can be expanded easily.
Methodology

- **MSD-cell: Generating Multi-Scale Description**

  the module needs to implement two key functions

  **First function: extract multi-scale description**
  - The first function can remember prior features on different scales and add new scale feature.
  - Implemented by a **1x3x32/1 convolutions**

  **Second function: mix multi-scale description**
  - the second function can compare and mix these two kinds of features.
  - Implemented by a **1x1x32/1 convolutions**

Fig. 3: The detailed structure of MSD-cell, which can be expanded easily.
**Methodology**

- **MSD-cell: Generating Multi-Scale Description**
  - The module needs to implement two key functions
  - **First function: extract multi-scale description**
    - The first function can remember prior features on different scales and add new scale feature.
    - Implemented by a $1x3x32/1$ convolutions
  - **Second function: mix multi-scale description**
    - The second function can compare and mix these two kinds of features.
    - Implemented by a $1x1x32/1$ convolutions

*Fig. 3: The detailed structure of MSD-cell, which can be expanded easily.*
**Methodology**

- **MSD-cell: Generating Multi-Scale Description**
  
  The module needs to implement two key functions.

  **First function: extract multi-scale description**
  
  - The first function can remember prior features on different scales and add new scale feature.
  - Implemented by a `1x3x32/1` convolutions

  **Second function: mix multi-scale description**
  
  - The second function can compare and mix these two kinds of features.
  - Implemented by a `1x1x32/1` convolutions

Fig. 3: The detailed structure of MSD-cell, which can be expanded easily.
Methodology

- MSD-cell: Generating Multi-Scale Description
  
  The module needs to implement two key functions:

  First function: extract multi-scale description
  - The first function can remember prior features on different scales and add new scale features.
  - Implemented by a 1x3x32/1 convolutions

  Second function: mix multi-scale description
  - The second function can compare and mix these two kinds of features.
  - Implemented by a 1x1x32/1 convolutions

Fig. 3: The detailed structure of MSD-cell, which can be expanded easily.
**Methodology**

- **MSD-cell: Generating Multi-Scale Description**
  
  The module needs to implement two key functions:
  
  **First function: extract multi-scale description**
  - The first function can remember prior features on different scales and add new scale feature.
  - Implemented by a $1 \times 3 \times 32 / 1$ convolutions

  **Second function: mix multi-scale description**
  - The second function can compare and mix these two kinds of features.
  - Implemented by a $1 \times 1 \times 32 / 1$ convolutions

Fig. 3: The detailed structure of MSD-cell, which can be expanded easily.
Methodology

- **MSD-cell: Generating Multi-Scale Description**
  
  The module needs to implement two key functions:

  **First function: extract multi-scale description**
  - The first function can remember prior features on different scales and add new scale feature.
  - Implemented by a **1x3x32/1 convolutions**

  **Second function: mix multi-scale description**
  - The second function can compare and mix these two kinds of features.
  - Implemented by a **1x1x32/1 convolutions**

Fig. 3: The detailed structure of MSD-cell, which can be expanded easily.
Methodology

◆ MSDNN

Multi-Scale Description based Neural Network

Shared part
◆ The shared part can extract multi-scale description feature from seismic waveforms with three spatial dimensions.
◆ Implemented by 1 \text{conv}(1 \times 3) \text{ and } 10 \text{MSD-cell}

Detection part
◆ the Detection part can distinguish multi-scale description features for earthquake detection.
◆ Implemented by \text{fc}(128), \text{fc}(2) \text{ and softmax}

Fig. 4: The framework of MSDNN, which is divided into shared part and detection part.
Methodology

- **MSDNN**
  
  Multi-Scale Description based Neural Network

**Shared part**
- The shared part can extract multi-scale description feature from seismic waveforms with three spatial dimensions.
- Implemented by 1 conv(1x3) and 10 MSD-cell

**Detection part**
- The Detection part can distinguish multi-scale description features for earthquake detection.
- Implemented by fc(128), fc(2) and softmax

Fig. 4: The framework of MSDNN, which is divided into shared part and detection part.
Homologous Earthquake Waveforms

- Homologous Earthquake waveforms are one earthquake detected by multiple monitoring stations.
- Each pair of waveforms can be labeled as positive pair (means homology) or negative pair (means non-homology)

Fig. 5: Sampling pairs of multi-task learning.
Methodology

- **Homologous Earthquake Waveforms**
  - Homologous Earthquake waveforms are one earthquake detected by multiple monitoring stations.
  - Each pair of waveforms can be labeled as positive pair (means homology) or negative pair (means non-homology)

Fig. 5: Sampling pairs of multi-task learning.
Methodology

Homologous Earthquake Waveforms

- Homologous Earthquake waveforms are one earthquake detected by multiple monitoring stations.
- Each pair of waveforms can be labeled as positive pair (means homology) or negative pair (means non-homology)

Fig. 5: Sampling pairs of multi-task learning.
Methodology

- **Multi-Task Learning Strategy**

  - **Auxiliary task**: homologous earthquake detection task.
  - **Main task**: aftershock detection task.

  - Auxiliary task can optimize the multi-scale description feature

Fig. 4: The framework of MSDNN and multi-task learning, which is divided into shared, detection and auxiliary part.
Methodology

- **Multi-Task Learning Strategy**
  - **Auxiliary task**: homologous earthquake detection task.
  - **Main task**: aftershock detection task.
  - Auxiliary task can optimize the multi-scale description feature.

Fig. 4: The framework of MSDNN and multi-task learning, which is divided into shared, detection and auxiliary part.
Methodology

- **Multi-Task Learning Strategy**
  - **Auxiliary task**: homologous earthquake detection task.
  - **Main task**: aftershock detection task.
  - **Auxiliary task** can optimize the multi-scale description feature.

Fig. 4: The framework of MSDNN and multi-task learning, which is divided into shared, detection and auxiliary part.
Experiments

- **Overall Performance**
  - our **MSDNN** methods consistently outperform all the baselines.
  - the **Multi-task Learning** can improve performance in all metrics.

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy</th>
<th>Recall</th>
<th>Precision</th>
<th>$F_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic Regression</td>
<td>0.505</td>
<td>0.520</td>
<td>0.080</td>
<td>0.130</td>
</tr>
<tr>
<td>Support Vector Machine</td>
<td>0.515</td>
<td>0.520</td>
<td>0.080</td>
<td>0.130</td>
</tr>
<tr>
<td>Random Forest [40]</td>
<td>0.767</td>
<td>0.680</td>
<td>0.190</td>
<td>0.300</td>
</tr>
<tr>
<td>XGboost [41]</td>
<td>0.882</td>
<td>0.770</td>
<td>0.350</td>
<td>0.490</td>
</tr>
<tr>
<td>ConvNetQuake [7]</td>
<td>0.935</td>
<td>0.602</td>
<td>0.544</td>
<td>0.571</td>
</tr>
<tr>
<td>Inception Net [11]</td>
<td>0.941</td>
<td>0.637</td>
<td>0.582</td>
<td>0.608</td>
</tr>
</tbody>
</table>

**Our Solutions**

<table>
<thead>
<tr>
<th></th>
<th>Accuracy</th>
<th>Recall</th>
<th>Precision</th>
<th>$F_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSDNN</td>
<td>0.952</td>
<td>0.638</td>
<td>0.678</td>
<td>0.658</td>
</tr>
<tr>
<td>MSDNN+Multi-task Learning</td>
<td>0.954</td>
<td>0.667</td>
<td>0.683</td>
<td>0.675</td>
</tr>
</tbody>
</table>
Experiments

- **Overall Performance**
  - our MSDNN methods consistently outperform all the baselines.
  - the Multi-task Learning can improve performance in all metrics.

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy</th>
<th>Recall</th>
<th>Precision</th>
<th>$F_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic Regression</td>
<td>0.505</td>
<td>0.520</td>
<td>0.080</td>
<td>0.130</td>
</tr>
<tr>
<td>Support Vector Machine</td>
<td>0.515</td>
<td>0.520</td>
<td>0.080</td>
<td>0.130</td>
</tr>
<tr>
<td>Random Forest [40]</td>
<td>0.767</td>
<td>0.680</td>
<td>0.190</td>
<td>0.300</td>
</tr>
<tr>
<td>XGboost [41]</td>
<td>0.882</td>
<td>0.770</td>
<td>0.350</td>
<td>0.490</td>
</tr>
<tr>
<td>ConvNetQuake [7]</td>
<td>0.935</td>
<td>0.602</td>
<td>0.544</td>
<td>0.571</td>
</tr>
<tr>
<td>Inception Net [11]</td>
<td>0.941</td>
<td>0.637</td>
<td>0.582</td>
<td>0.608</td>
</tr>
<tr>
<td><strong>Our Solutions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSDNN</td>
<td>0.952</td>
<td>0.638</td>
<td>0.678</td>
<td>0.658</td>
</tr>
<tr>
<td>MSDNN+Multi-task Learning</td>
<td>0.954</td>
<td>0.667</td>
<td>0.683</td>
<td>0.675</td>
</tr>
</tbody>
</table>
Experiments

- **Overall Performance**
  - our **MSDNN** methods consistently outperform all the baselines.
  - the **Multi-task Learning** can improve performance in all metrics.

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy</th>
<th>Recall</th>
<th>Precision</th>
<th>$F_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic Regression</td>
<td>0.505</td>
<td>0.520</td>
<td>0.080</td>
<td>0.130</td>
</tr>
<tr>
<td>Support Vector Machine</td>
<td>0.515</td>
<td>0.520</td>
<td>0.080</td>
<td>0.130</td>
</tr>
<tr>
<td>Random Forest [40]</td>
<td>0.767</td>
<td>0.680</td>
<td>0.190</td>
<td>0.300</td>
</tr>
<tr>
<td>XGboost [41]</td>
<td>0.882</td>
<td>0.770</td>
<td>0.350</td>
<td>0.490</td>
</tr>
<tr>
<td>ConvNetQuake [7]</td>
<td>0.935</td>
<td>0.602</td>
<td>0.544</td>
<td>0.571</td>
</tr>
<tr>
<td>Inception Net [11]</td>
<td>0.941</td>
<td>0.637</td>
<td>0.582</td>
<td>0.608</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Our Solutions</strong></th>
<th>Accuracy</th>
<th>Recall</th>
<th>Precision</th>
<th>$F_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSDNN</td>
<td>0.952</td>
<td>0.638</td>
<td>0.678</td>
<td>0.658</td>
</tr>
<tr>
<td>MSDNN+Multi-task Learning</td>
<td>0.954</td>
<td>0.667</td>
<td>0.683</td>
<td>0.675</td>
</tr>
</tbody>
</table>
Experiments

◆ Case Studies

◆ The “False Positive” waveforms and the “True Positive” waveforms are similar.

◆ Several geophysical experts re-check 163 “False Positive” waveforms, and 161 waveforms of them were labeled as “Positive”. MSDNN framework can help manual checking.
Case Studies

- The “False Positive” waveforms and the “True Positive” waveforms are similar.
- Several geophysical experts re-check 163 “False Positive” waveforms, and 161 waveforms of them were labeled as “Positive”. MSDNN framework can help manual checking.
Experiments

- **Case Studies**
  - The “False Positive” waveforms and the “True Positive” waveforms are similar.
  - Several geophysical experts re-check 163 “False Positive” waveforms, and 161 waveforms of them were labeled as “Positive”. MSDNN framework can help manual checking.
Experiments

- **Case Studies**
  - The “False Positive” waveforms and the “True Positive” waveforms are similar.
  - Several geophysical experts re-check 163 “False Positive” waveforms, and 161 waveforms of them were labeled as “Positive”. MSDNN framework can help manual checking.
Experiments

*Case Studies*

- The “False Positive” waveforms and the “True Positive” waveforms are similar.
- Several geophysical experts re-check 163 “False Positive” waveforms, and 161 waveforms of them were labeled as “Positive”. MSDNN framework can help manual checking.

![Waveform Images]

(a) Average of False Positive
(b) True Positive
(c) False Positive
Experiments

- **Case Studies**
  - The “False Positive” waveforms and the “True Positive” waveforms are similar.
  - Several geophysical experts re-check 163 “False Positive” waveforms, and 161 waveforms of them were labeled as “Positive”. MSDNN framework can help manual checking.
Experiments

- **PCA Visualization**
  - Blue dots are the positive samples and the red dots are negative samples.
  - The positive blue dots are close together and easily separated from the red dots when multi-task learning is deployed.

(a) MSDNN  
(b) MSDNN+Multi-task Learning
Experiments

- **PCA Visualization**
  - Blue dots are the positive samples and the red dots are negative samples.
  - The positive blue dots are close together and easily separated from the red dots when multi-task learning is deployed.

![PCA Visualization](image)
Experiments

- **PCA Visualization**
  - Blue dots are the positive samples and the red dots are negative samples.
  - The **positive blue dots are close together** and easily separated from the red dots when multi-task learning is deployed.
Conclusion

◆ Contributions
  
 ◆ We propose a novel neural network based solution: **MSDNN (Multi-Scale Description based Neural Network)**

 ◆ We design a **multi-task learning** strategy for utilizing the relationship between different monitoring stations.

 ◆ We evaluate our framework on a **real-world data set** from aftershocks of the Wenchuan M8.0 Earthquake.

◆ Future Work
  
 ◆ Earthquake Rapid Report and Time-series Event Detection
Thank you for listening!

Q&A

Reporter: Qi Zhang

zq26@mail.ustc.edu.cn