

Deep Convolutional Neural Network with Bottleneck Structure using Raw Seismic Waveform for Earthquake Classification

Bon-Hwa Ku*, Gwan-Tae Kim**, Jeong-Ki Min***, Hanseok Ko***

Abstract

In this paper, we propose deep convolutional neural network(CNN) with bottleneck structure which improves the performance of earthquake classification. In order to address all possible forms of earthquakes including micro-earthquakes and artificial-earthquakes as well as large earthquakes, we need a representation and classifier that can effectively discriminate seismic waveforms in adverse conditions. In particular, to robustly classify seismic waveforms even in low snr, a deep CNN with 1x1 convolution bottleneck structure is proposed in raw seismic waveforms. The representative experimental results show that the proposed method is effective for noisy seismic waveforms and outperforms the previous state-of-the art methods on domestic earthquake database.

▶Keyword: Convolutional neural network, earthquake classification, bottleneck structure, raw seismic waveform, centering preprocessing

1. Introduction

시계열 데이터(Time series data)는 음성 인식, 금융 분석, 지진 분석, 생체 신호 분석 등 다양한 분야에서 발생한다. 주식시장의 주가 예측이나 심전도를 이용한 심장병 예측과 같이 시계열 예측 문제는 일정 간격으로 취득한 데이터의 수열을 기반으로 예측 모델을 생성한 후 미래에 일어날 사건들을 추정하게 된다. 시계열 문제의 일반적인 접근법은 특징 추출과 기계 학습으로 대표된다. 전통적인 특징 기반 시계열 문제에서는 특징 추출과 학습기의 파라미터들을 전문가의 경험적/실험적 요소에 따라 설정하게 된다. 최근에 영상 분류에서 CNN(Convolutional Neural Network)으로 대표되는 딥러닝(Deep learning)의 성공은 다른 분야로도 전파되고 있다. 시계열 문제에서도 기존의 전통적인 방식에서 벗어나 딥러닝을 이용하여 효율적인 특징 추출 및 기계 학습을 하는 방식으로 전환되고 있다.

본 논문에서는 대표적인 시계열 데이터인 지진파를 이용한

지진 이벤트 분류 문제를 다룬다. 기존의 전통적인 방식에서 벗어나 다양한 딥러닝을 적용한 접근 방식들이 제시되고 있으며 특히 CNN 기반 지진 이벤트 분류 연구에서는 컨볼루션(Convolution)과 비선형 변환 과정을 기반으로 지진 이벤트에 적합한 특징 추출 및 분류를 하였다. 본 논문에서는 기존 방식을 확장하여 병목(Bottleneck) 구조가 적용된 CNN을 이용한 지진 이벤트 분류 기법을 제안한다. 병목 구조는 최근 GoogLeNet, ResNet 등 다양한 딥러닝 기법에서 연산량 및 파라미터 감소를 위해 적용되었다. 본 논문에서는 병목 구조의 차원 축소(Dimension reduction)에서 파생되는 잡음제거 효과를 통해 지진 이벤트 분류 성능을 향상시키고자 한다. 본 논문에서 다루는 지진 이벤트는 강진, 미소 지진, 인공 지진, 잠음으로 나뉜다. 수집된 강진, 미소지진, 인공지진의 경우 지진 이벤트 정보 뿐만 아니라 환경 및 센서 잡음 요소도 같이 기록되게 된다.

• First Author: Bon-Hwa Ku, Corresponding Author: Hanseok Ko

*Bon-Hwa Ku (bhku@ispl.korea.ac.kr), Machine learning and Big Data Research Center, Korea University

**Gwan-Tae Kim (gtkim@ispl.korea.ac.kr), Dept. of Visual Information Processing, Korea University

***Jeong-Ki Min (jkmin@ispl.korea.ac.kr), School of Electrical Engineering, Korea University

****Hanseok Ko (hsko@korea.ac.kr), School of Electrical Engineering, Korea University

• Received: 2018. 12. 06, Revised: 2019. 01. 07, Accepted: 2019. 01. 08.

• This work was supported by Meteorological/Earthquake See-At Technology Development Research Grant (KMI2018-09610).

1x1 컨볼루션을 이용한 병목 구조는 주성분 분석(Principle Component Analysis)에서의 차원 축소 및 복원 과정에서의 잡음 제거 효과를 모델 훈련 과정에서 획득하고자 한다. 또한 본 논문에서는 입력 분포가 상이해지는 내부 공분산 이동(Internal covariance shift) 현상을 방지하는 일괄 정규화(Batch normalization)과 dropout 정규화(Regularization)를 통해 학습 모델의 안정성 및 성능 향상을 도모한다. 일괄 정규화 및 dropout 정규화 기법을 지진 이벤트 분류에 적합하게 배치한 CNN 구조를 제시한다.

본 논문의 구성은 다음과 같다. 3장에서는 제안된 CNN을 이용한 지진 이벤트 분류 기법을 제시 하였으며 4장에서는 학습 및 테스트에 이용하는 지진 데이터의 수집에 관해 설명하였다. 5장에서는 제안된 알고리즘의 성능 평가를 다루며, 6장에서 결론을 맺는다.

II. Preliminaries

1. Related works

시계열 데이터의 이벤트 분석 및 분류 문제에서도 딥러닝을 이용한 방식들이 최근 보고 되고 있다. [1]에서는 다양한 스케일에서의 시계열 데이터 문제를 풀기 위해 다중 스케일 CNN(Multi-scale Convolutional Neural Network) 기법을 제안하였다. 다중 스케일 CNN은 특징 추출 및 분류를 단일 프레임 워크에서 통합된 구조를 갖고 있으며 다양한 스케일과 주파수 특징을 자동으로 추출한다. [2]에서는 음악 장르 분류 문제에서 미가공 웨이브(Raw waveform) 샘플 데이터에 기반한 CNN 모델을 제안하였다. 기존의 프레임 레벨(Frame level) 스펙트로그램(Spectrogram) 모델에서는 페이즈(Phase) 정보가 사라지거나 시간-주파수 해상도 균형점(time frequency resolution trade off) 문제가 발생하는 반면 제안된 방법에서는 미가공 데이터를 입력으로 작은 컨볼루션 필터를 깊은 구조로 적용하여 프레임 레벨 스펙트로그램 모델에서 발생하는 문제들을 극복하고자 하였다. 의공학 분야에서는 심전도 신호를 기반으로 생체 인증을 위한 웨이블릿 공간(Wavelet domain) 다중 해상도 CNN 모델이 제안되었다[3]. 데이터 표현의 다양성을 위해 무작위로 선택한 샘플들을 웨이블릿 공간으로 변환하여 다중 해상도 시간-주파수 공간에서 표현한 후 다중 해상도 1-D CNN을 적용한다.

지진 탐지 혹은 지진 이벤트 분류는 시계열 분석의 대표적인 분야로 STA/LTA 방식은 STA(Short Time Average)와 LTA(Long Time Average)의 비율 문턱치를 통해 지진을 탐지하는 대표적인 방식이다. 이 방식은 강진 탐지 및 실용성 측면에서 우수한 면을 보이지만 낮은 SNR 환경에서는 부적합하다 [4]. 자기상관(Autocorrelation) 기반 방식은 단일 지역에서 반복되는 지진 파형들간의 상관 분석(Correlation analysis)을 기반으로 지진을 탐지하는 방식으로 성능은 우수하지만 연산량

과부하 및 긴 시간의 신호에는 활용성에 한계를 갖고 있다[5]. 템플릿 매칭(Template matching) 방식은 자기상관 방법의 연산량 과부하 문제를 개선하기 위해 제안된 기법으로 지진 템플릿을 설정 후 템플릿과 입력 지진 파형과의 상관 분석을 통해 지진을 탐지하게 된다[6]. 템플릿의 숫자는 성능에 주요한 영향을 미치며 주성분 분석 혹은 부공간 학습(Subspace learning)을 통해 템플릿의 숫자를 줄이는 접근 방식들이 제안되고 있다. 최근에는 템플릿 매칭의 연산량을 개선하기 위한 FAST(Fingerprint And Similarity Thresholding) 방법이 제안되었다[7]. 이 방식은 비지도학습(Unsupervised learning) 지진 탐지 방식으로 스펙트럼 영상(Spectral image)과 Haar 웨이블릿(Harr wavelet) 변환을 통해 지문과 유사한 지진 특징을 추출하며 국소 local sensitive hashing을 통해 유사도 측정 연산량을 감소시킨다. 최근 딥러닝을 적용한 지진 탐지 기법도 다양하게 보고되고 있다. 특히 [8]에서는 3 채널의 미가공 지진과 데이터를 이용하여 CNN 모델을 학습하는 방식을 제안하였다. 제안된 방식은 시계열 분류 문제에서의 딥러닝 모델과 유사한 형태를 갖고 있으며 총 8 레이어의 1D CNN과 비선형 변환 Relu 과정을 수행한다.

III. Proposed Scheme

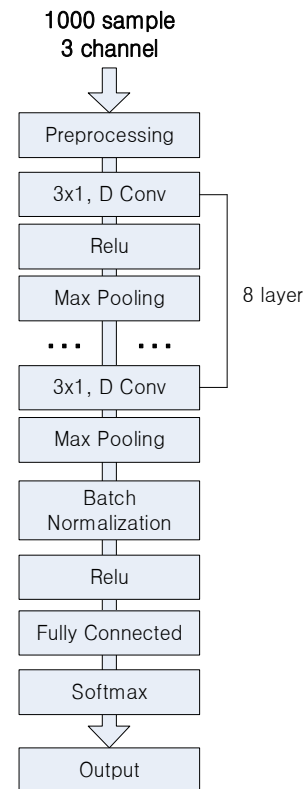


Fig. 1. Earthquake classifier using CNN

시계열 데이터를 기반으로 한 딥러닝 기법들은 크게 스펙트로그램(Spectrogram)과 같은 2차원 변환 입력을 이용하는 방식과 시계열 원시 데이터(Time series raw data)를 바로 이용하는 방식으로 분류할 수 있다. 지진과 분류의 경우 주파수 크기뿐만 아니라 위상 천이 정보가 주요 정보로 적용되어야 한다. 그러나 스펙트로그램과 같은 2차원 변환 입력 방식의 경우 입력 데이터의 위상 변화 정보를 사용하지 않기 때문에 본 논문에서는 그림 1과 같은 시계열 원시 데이터를 이용하는 구조의 CNN 방법을 적용하였다.

3.1 Preprocessing

관측소별 특성으로 인해 지진 데이터의 범위가 다른 경우들이 발생하는 경우 식 (1)과 중심화(Centering) 과정의 데이터 전처리 과정을 수행한다. 일반적으로 중심화 과정은 기하학적으로 데이터의 균집을 모든 차원에 대해서 원점으로 이동시키는 효과를 얻을 수 있다.

$$y_m = x_m - \frac{1}{M} \sum_m^M x_m \quad (1)$$

여기서 x_m 은 지진 샘플데이터, y_m 은 센터링 변환된 지진 샘플 데이터, M은 발생 이벤트내의 총 샘플수를 나타낸다.

3.2 Network structure

제안하는 CNN 구조는 3채널 지진 데이터를 이용하며 채널당 1000 샘플의 입력 데이터를 사용한다. 1000 샘플 입력 데이터는 100Hz 샘플링 10초 구간의 데이터를 나타낸다. 8개의 컨볼루션 레이어(Convolution layer)은 식 (2)와 같이 1D 컨볼루션을 수행하며 컨볼루션 결과는 식(3)과 같은 Relu(Rectified Linear Unit) 함수를 통해 비선형 변환을 거친다.

$$x_{c,t}^l = R \left(\sum_c^C \sum_t^3 x_{c,t+t'-1}^{l-1} \cdot W_{c,c,t}^l \right) \quad \text{for } l=1, \dots, 8 \quad (2)$$

$$R(z) = \max(0, z) \quad (3)$$

여기서 x^{l-1} 은 입력 지진 샘플데이터, W^l 은 필터, C는 채널의 수를 나타낸다. 1D 컨볼루션의 필터 크기는 3x1 형태를 사용하며 채널 C의 크기는 각 레이어마다 32개를 적용한다(단, 첫 번째 레이어에서 채널의 크기는 3으로 설정한다). 변환 결과는 max-pooling을 통해 downsampling 과정을 거치게 된다. Max-pooling의 간격은 2로 설정하여 이 과정을 거칠 때 마다 입력 데이터는 1/2배로 축소된다. 8 레이어의 컨볼루션 후에 생성된 특징 정보(Feature map)는 일괄 정규화를 통해 정규화된 특징 정보를 생성하며 한번 더 비선형 Relu 과정을 수행한다. 생성된 특징 정보는 fully-connected 레이어와 softmax를 통해 지진 이벤트 분류를 수행한다. 제안 모델의 과대적합(overfitting)을 방지하기 위해 본 논문에서는 fully-connected

레이어에 dropout 정규화를 적용하였다.

3.3 Batch Normalization

네트워크 레이어(Network layer)가 증가하면서 경사도 사라짐(Gradient vanishing) 및 폭발(Exploding) 문제는 딥네트워크(deep network)의 주요 쟁점 사항이다. 이러한 문제는 활성화(Activation) 함수의 변화, 초기화(Initialization), 학습률(Learning rate)을 통해 해결하려는 다양한 시도가 있었다. 일괄 정규화는 경사도 사라짐 및 폭발의 문제를 훈련(Training) 과정 안에서 해결하려는 접근을 취한다. 일괄 정규화는[9] 각 레이어마다 입력 분포가 상이해지는 내부 공분산 이동 현상을 막기 위해 각 레이어의 입력 분포를 평균 0, 표준편차 1이 형태로 정규화시키는 whitening 방법을 착안하여 고안되었다. 각 특징에 대해서 평균과 분산을 구한 후 정규화를 거치며 정규화된 값들에 대해서 스케일 요소(scale factor)와 이동 요소(shift factor)를 더해 주는 변환 과정을 거치게 된다. 스케일 요소와 이동 요소 파라미터는 mini-batch 단위로 훈련 과정에서 유도한다.

3.4 Dropout regularization

Srivastava[10]에 의해 제안된 dropout 기법은 과대적합을 방지하기 위한 정규화 기법이다. dropout은 각 뉴런들을 확률 값에 따라 활성화하여 학습에 적용하게 된다. 본 논문에서는 fully connected layer에서 p = 0.5를 설정하여 네트워크를 훈련한다.

3.5 CNN with Bottleneck structure

병목 구조는 그림 2와 같이 각 레이어의 노드(Node) 수가 줄어들었다가 다시 증가하는 형태를 나타낸다. 이러한 구조는 특징의 차원 축소 후 다시 고차원으로 복원되는 과정으로 설명될 수 있다. 1x1 컨볼루션을 이용한 병목 구조는 GoogLeNet, ResNet에서 연산량 감소 및 파라미터가 증가하는 것을 방지한다.

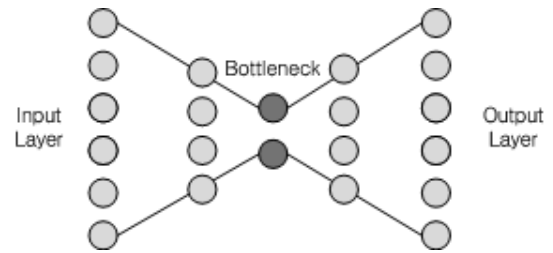


Fig. 2. Bottleneck structure

본 논문에서는 연산량 혹은 파라미터 감소 측면 보다는 잡음 제거 효과를 얻기 위해 1x1 컨볼루션을 이용한 병목 구조를 적용한다. 영상 분류 문제 데이터와는 다르게 지진 이벤트 데이터들은 실측 데이터들을 기반으로 학습 과정을 거쳐야 한다. 즉 지진 이벤트 내에는 지진 이외의 잡음 성분들도 포함될 가능성이 있다. 본 논문에서는 시계열 데이터 CNN에서 1D 컨볼루션 레이어를 그림 3과 같이 1x1, 3x3, 1x1 필터로 대체한 CNN 구조를 적용하였다. 이러한 구조는 주성분 분석 혹은 오토인코

더(Autoencoder)와 같이 차원 축소와 복원 과정을 통해 데이터내의 불필요한 잡음들을 제거하는 효과를 갖게 된다. 1x1 컨볼루션을 이용한 병목 구조는 레이어2부터 레이어8에 적용한다. 레이어 1에서는 컨볼루션 후 일괄 정규화를 수행한다.

IV. Dataset

본 논문에서는 2018년 1월1일부터 2018년 7월 31일까지의 대한민국에서 발생하는 지진 데이터 정보들을 이용하여 지진 데이터 및 잡음 데이터를 수집하였다. 지진 관측은 약 200여 곳의 관측소에서 관측을 하고 있으며 24시간 연속 지진 데이터들이 저장되어 있다. 그림 4와 같이 기상청에서 제공되는 24시간 지진 관측 데이터 및 지진 이벤트 목록 파일을 토대로 지진 이벤트 및 잡음 데이터베이스를 구축하였다. 그림 4의 상단은 24시간 관측 데이터를 나타내며 적색 구간에서 지진 이벤트가 발생하였다. 발생한 지진 이벤트는 지진 발생 시간을 기준으로 하거나 혹은 관측소 관측 시간을 기준으로 데이터가 추출된다. 그림 4의 하단 그림은 지진 이벤트 발생 시간을 기준으로 추출된 지진 이벤트 데이터이다. 본 실험에서는 관측소 관측 시간을 기준으로 10초간의 지진 이벤트 데이터를 수집하였다. 지진 관측 데이터는 100샘플 데이터만을 이용하고 있으며 강진(규모 2.0 이상), 미소지진(규모 2.0 미만), 인공지진, 잡음 카테고리 데이터베이스를 구성하였다. 수집 기간에 발생한 지진 이벤트의 횟수 및 각 지진 이벤트를 관측한 관측소의 개수는 표 1과 같다.

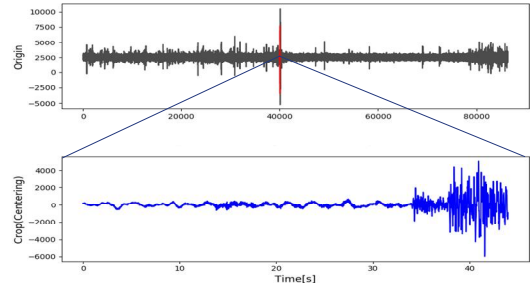


Fig. 4. Extract earthquake events from 24 hour continuous data

Table 1. Number of Earthquake events

Earthquake Event	Number of Occurrence	Number of Observatory
Macro-earthquake	71	250
Micro-earthquake	189	671
Man-made earthquake	73	179

본 논문에서 훈련 및 테스트에 이용될 지진 이벤트의 개수는 관측소 개수와 동일하게 된다. 수집 데이터의 양을 늘리기 위해 본 논문에서는 관측소에서 관측한 발생 시점 보다 3초 앞부터 10초 구간의 데이터를 수집하고 있으며 1초 간격 sliding 방식을 적용하여 데이터의 양을 4배 증가시켰다. 잡음 이벤트의 경우 지진 이벤트 목록에 기록되지 않은 날짜에서 임의적으로 10초 구간의 샘플을 2,469개 추출하였다. 모의 실험에서 사용하는 데이터셋 구성은 표 2와 같다. 데이터 균형을 맞추기 위해 미소 이벤트의 개수는 각 데이터셋 별로 상이하다.

Table 2. Event Dataset

Dataset	Total Number	Details
Earthquake(E) vs Noise	6,357	3,888 vs 2,469
Macro-E vs Noise	2,576	876 vs 1,700
Micro-E vs Noise	4,809	2,340 vs 2,469
Manmade-E vs Noise	2,072	672 vs 1,400
Micro-E vs Manmade-E	1,376	704 vs 672

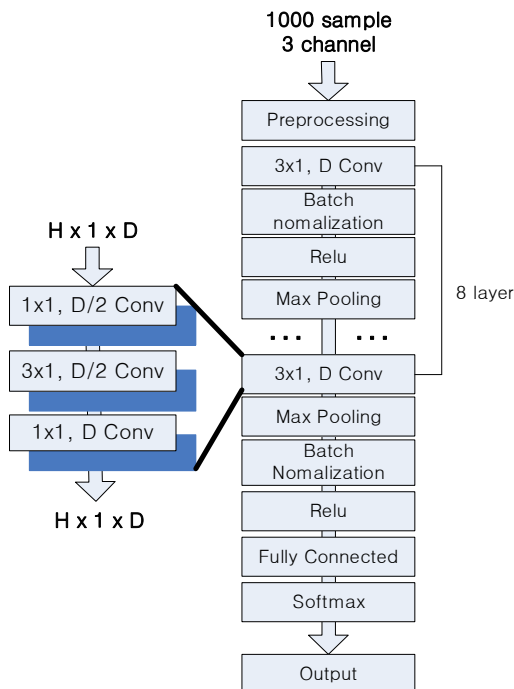


Fig. 3. CNN with 1x1 bottleneck structure

V. Experiments

5.1. Learning parameter and method

모의 실험에서는 표 2의 데이터셋에서 훈련 70%, validation 15%, 테스트 15%의 비율로 데이터를 사용하였다. NVIDIA Geforce GTX 1080Ti GPU와 Tensorflow 기반으로 모의 실험을 수행하였으며 네트워크 훈련 iteration은 총 30,000을 수행하였다. 모델 훈련시 최적화는 ADAM 방법을 이용하였으며 학습률은 10^{-4} 을 주었다.

5.2 Performance evaluation

제안한 방법의 성능 비교를 위해 ConvNetQuake[8]와 주파

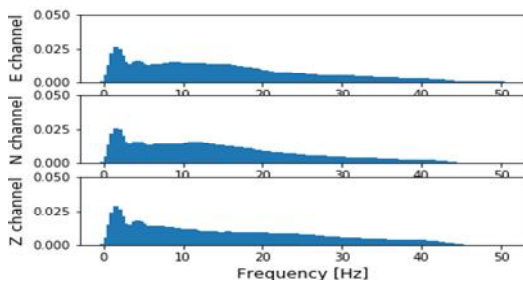
수 특징 + SVM 방법을 선택하였다. 성능 측정은 식(4)와 같이 정확도(Accuracy) 지수를 사용한다. 정확도는 총 테스트 개수 중에서 정확하게 추정된 결과(TP:true positive, TN:true negative) 개수의 비로 정의된다.

$$Accuracy = \frac{TP + TN}{Test\ Sample\ Number} \quad (4)$$

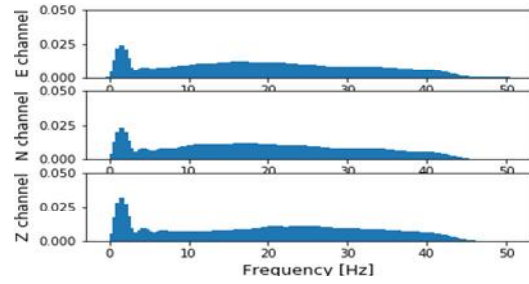
ConvNetQuake는 모의 실험데이터를 이용하여 훈련 및 테스트를 진행하였으며 모델 훈련 파라미터는 원저자가 설정한 파라미터와 동일하게 1D 컨볼루션 필터 크기는 3x1, 채널 32로 설정하였으며 downsampling 효과를 위해 필터 stride는 2 설정하였다. 이와 더불어 L_2 정규화 파라미터 λ 는 10^{-3} , 최적화 방법 및 학습율은 ADAM 및 10^{-4} 로 설정하였다.

SVM 방식에서는 129차원의 주파수 특징을 이용하여 지진 이벤트 분류기를 훈련하였다. 모델 훈련시 커널은 RBF 커널을 이용하였다. 이벤트별 주파수 히스토그램(Histogram)은 그림 5와 같다. 모든 이벤트에서 공통적으로 저주파 성분이 강하게 나타나고 있으며 인공지진에 비해 자연지진(강진, 미소지진)의 히스토그램 분포는 잡음과 유사한 형태를 보이고 있다.

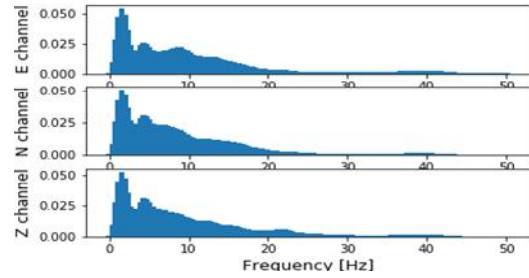
제안한 방법에서 1x1 병목 구조의 효용성을 알아보기 위해 모의 실험에서는 제안 구조를 (1), (2)로 구분하였다. 제안 방법 (1)은(그림 1) 1x1 병목 구조가 적용되지 않은 구조, 제안 방법 (2)는(그림 3) 1x1 병목 구조가 적용된 구조이다. 제안 방법의 성능 비교 결과는 표 3과 같다. 기존 CNN 방법의 경우 각 데이터셋 별로 80% 초중반대의 성능을 나타내고 있다. 기존의 handcraft 방식인 SVM 에 비해서는 월등히 우수한 성능 향상을 보여주고 있다. 제안한 방법(1)은 ConvNetQuake 방법보다 평균적으로 10.2 % 향상된 성능을 보여주고 있다. 일괄정규화와 dropout을 적용함으로써 L_2 정규화만이 적용된 ConvNetQuake 보다 우수한 성능을 얻을 수 있었다. 또한 제안 방법(2)는 ConvNetQuake 보다는 11.6%, 제안 방법 (1) 보다는 1.4% 성능 향상을 보이고 있다. 제안된 방법 (2)는 요소 이벤트(Sub-event)들로 구성된 지진과 잡음간의 분류 성능에 더욱 더 향상된 효과를 얻을 수 있었다.



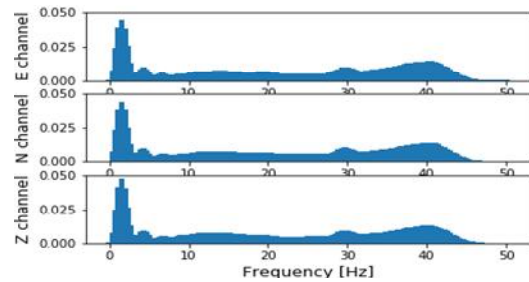
(a) Macro-earthquake



(b) Micro-earthquake



(c) Man-made earthquake



(d) Noise

Fig. 5. Frequency histogram distribution according events

Table 3. Comparison performance of proposed method (unit: %)

	SVM	ConvNet Quake[8]	Proposed CNN(1)	Proposed CNN(2)
Earthquake(E) vs Noise	62	84	90	93
Macro-E vs Noise	76	84	94	95
Micro-E vs Noise	68	85	92	93
Manmade-E vs Noise	79	81	95	96
Micro-E vs Manmade-E	57	81	95	96

VI. Conclusions

본 논문에서는 지진 분류 성능 향상을 위해 병목 구조의 시계열 CNN 기반 지진 분류 기법을 제안하였다. 낮은 SNR 환경에서도 강인하게 지진파를 분류하기 위해 1x1 컨볼루션 병목 구조를 적용한 CNN 기반 지진 분류 기법이 제안되었으며 제안

된 기법은 기존의 SVM 방식 및 기존 CNN을 이용한 방식보다 평균 26.2%, 11.6% 향상된 성능을 얻을 수가 있었다. 일괄 정규화 및 dropout 정규화를 통해 기존 CNN 구조의 지진 분류 기법보다 향상된 결과를 얻을 수 있었으며 1x1 컨볼루션 구조를 통해 잡음과의 구분성을 향상시킴으로써 지진/잡음 분류 성능을 향상 시킬 수 있었다. 현재 인공 지진의 데이터는 다른 이벤트에 비해 충분히 확보되지 않아 모든 카테고리(강진/미소/인공/잡음)에 대한 성능평가는 수행하지 않았다. 추후 인공 지진의 데이터베이스를 확보 후 제한한 알고리즘의 모든 카테고리에 대한 분류 성능도 살펴볼 예정이다.

REFERENCES

- [1] Zhicheng Cui, Wenlin Chen, Yixin Chen, "Multi-Scale Convolutional Neural Networks for Time Series Classification," arXiv:1603.06995, May 2016.
- [2] Jogpil Lee, et al., "SampleCNN : End-to-End Deep Convolutional Neural Networks Using Very Small Filter for Music Classifications," Applied Science, Vol.8, No.1, pp.150, Jan. 2018.
- [3] Qingxue Zhang, et al., "HeartID: a multiresolution convolutional neural network for ECG-based biometric human identification in smart health applications," IEEE Access, Vol.5, pp.11805-11816, May 2017.
- [4] M. Withers, et al., "A comparison of select trigger algorithms for automated global seismic phase and event detection," Bull. Seismol. Soc. Am. Vol.88, No.1, pp.95-106, Feb. 1998.
- [5] S. J. Gibbons and F. Ringdal, "The detection of low magnitude seismic events using array-based waveform correlation," Geophys. J. Int. Vol.165, No.1 pp.149-166, Mar. 2006.
- [6] R. J. Skoumal, et al., "Optimizing multi-station earthquake template matching through re-examination of the Youngstown, Ohio, sequence," Earth and Planet Science Letters, Vol.405, pp.274-280, Nov. 2014.
- [7] C.E. Yoon, et al., "Earthquake detection through computationally efficient similarity search," Science Advances, Vol.1, No.11, e1501057, Dec. 2015.
- [8] Thibaut Perol, et al., "Convolutional neural network for earthquake detection and location," Science Advances, Vol. 4, No. 2, e1700578, Feb. 2018.
- [9] Sergey Ioffe and Christian Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," arXiv:1502.03167, Mar. 2015.
- [10] Nitish Srivastava, et al., "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," Journal of Machine Learning Research, Vol. 15, No.1, pp.1929-1958, Jun. 2014.
- [11] Dieleman, S. and Schrauwen, B., "End-to-end learning for music audio," In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, pp.6964-6968, 2014.
- [12] Ardila, D., et al., "Audio deepdream: Optimizing raw audio with convolutional networks," In Proceedings of the International Society for Music Information Retrieval Conference, 2016.
- [13] Thickstun, J., et al., "Learning features of music from scratch." arXiv, Apr. 2017.
- [14] Y. LeCun, Y. Bengio and G. Hinton, "Deep learning," Nature, Vol.521, No.7553, pp. 436-444, May 2015.
- [15] X. Cui, V. Goel and B. Kingsbury, "Data augmentation for deep neural network acoustic modeling," IEEE/ACM Trans. on Audio, Speech and Language Processing, Vol.23, No. 9, pp.1469-1477, Jun. 2015.
- [16] Simonyan, K. and Zisserman, A, "Very deep convolutional networks for large-scale image recognition," arXiv:1409.1556, Apr. 2015.
- [17] Anastasia Borovykh, et al., "Conditional Time Series Forecasting with Convolutional Neural Networks," arXiv:1703.04691, Sep. 2018.
- [18] Nima Hatami, et al., "Classification of Time-Series Images Using Deep Convolutional Neural Networks," arXiv:1710.00886, Oct. 2017.
- [20] Hassan Ismail Fawaz, et al., "Deep learning for time series classification: a review," arXiv: 1809.04356, Sep 2018.
- [21] NECIS, <http://necis.kma.go.kr/>

Authors



Bon-Hwa Ku received B.S. degree in Control Instrument Engineering from Kwangwoon University in 2000, M.S. degree in Electronics Engineering from Korea University, in 2002 and Ph.D. degree in Visual Information Processing Engineering

from Korea University, Seoul, in 2008. Since 2009, he has been with Korea University where he is currently Research Assistant Professor. His current research interests include machine learning and data fusion in computer vision.



Gwan-Tae Kim received the B.S. degrees in Mechanical Engineering from Korea Aerospace University, Korea, in 2018. He is currently Master's and Doctoral Integration Course Student in the Department of Visual Information

Processing, Korea University. He is interested in machine learning, reinforcement learning and intelligent agent.



Jeong-Ki Min is received the B.S. degree in Computer Engineering from Korea Polytechnic University. He is now working toward a M.S degree at Intelligent Signal Processing Laboratory, the School of Electrical Engineering, Korea University,

Korea. His current research interests include object detection, object tracking and machine learning.



Hanseok Ko received B.S. degree from Carnegie Mellon University, in 1982, M.S. degree from the Johns Hopkins University, in 1988, and Ph.D. degree from the CUA, in 1992, all in electrical engineering. At the onset of his career, he was with the WOL,

Maryland, where his work involved signal and image processing. He is Professor of Electrical and Computer Engineering and Director of Machine Learning and Big Data Institute at Korea University. He joined the faculty of ECE, Korea University, in 1995. He is currently serving as Guest-Editor for Sensors Journal on the special issue addressing multisensor fusion strategies. He is a Fellow of IET with his research interest in signal processing and machine learning for video analytics and human-machine interface.