

# 심층 신경망을 활용한 손상된 음성파일 복원 자동화

## Restoration of damaged speech files using deep neural networks

허희수,<sup>1</sup> 소병민,<sup>2</sup> 양일호,<sup>1</sup> 윤성현,<sup>1</sup> 유하진<sup>†</sup>

(Hee-Soo Heo,<sup>1</sup> Byung-Min So,<sup>2</sup> IL-Ho Yang,<sup>1</sup> Sung-Hyun Yoon,<sup>1</sup> and Ha-Jin Yu<sup>1†</sup>)

<sup>1</sup>서울시립대학교 컴퓨터과학부, <sup>2</sup>대검찰청

(Received December 8, 2016; revised January 11, 2017; accepted March 29, 2017)

**초 록:** 본 논문에서는 심층 신경망을 이용하여 손상된 음성파일을 복원하는 방법을 제안한다. 본 논문에서 다루는 음성파일 복원은 기존의 파일 카빙(file carving) 기반의 복원과는 다른 개념으로, 기존 기법으로는 복원할 수 없었던 손실된 정보를 복원하는 것이 목적이다. 새로운 복원 기법을 수행하는 과정에서 필요한 작업이지만 사람이 직접 수행할 수 없거나 너무 많은 시간이 소요되는 작업을 심층 신경망을 활용해 자동화할 수 있는 방안을 제안하였으며 관련한 실험을 진행하였다. 실험 결과, 심층 신경망을 활용해 음성, 비음성 분류나 음성파일 부호화 방식의 식별이 가능해 기존 파일 카빙 기반 방법이 복원하지 못하는 파일을 복원할 수 있었다.

**핵심용어:** 컴퓨터 법의학, 음성파일 복원, 심층 신경망, LSTM(Long Short Term Memory)

**ABSTRACT:** In this paper, we propose a method for restoring damaged audio files using deep neural network. It is different from the conventional file carving based restoration. The purpose of our method is to infer lost information which can not be restored by existing techniques such as the file carving. We have devised methods that can automate the tasks which are essential for the restoring but are inappropriate for humans. As a result of this study it has been shown that it is possible to restore the damaged files, which the conventional file carving method could not, by using tasks such as speech or nonspeech decision and speech encoder recognizer using a deep neural network.

**Keywords:** Digital forensic, Restoring audio file, Deep neural network, LSTM (Long Short Term Memory)

**PACS numbers:** 43.72.Uv, 43.60.Bf

### 1. 서 론

최근 스마트폰 등의 음성 녹음기기가 널리 보급되면서 음성파일을 법정 증거물로 제출하는 사례 또한 증가하고 있다. 법정 증거물로 제출된 음성파일은 조사관의 청구에 의해 그 내용이 확인되고 위변조 여부 검사 등을 통해 법정 증거물로서의 효력을 인정받게 된다. 하지만 스마트폰 등의 디지털 매체를 통해 수집된 음성파일은 악의적인 목적이나 메모리 부족 등의 이유로 삭제될 수 있다. 증거물로서 유

효한 내용을 담고 있는 음성파일이라고 하더라도 파일이 삭제된 경우에는 그 내용을 확인할 수 없기 때문에 삭제된 파일을 복원하는 과정이 필요하다. 삭제된 음성파일이 증거로서 인정받기 위해서는 음성파일이 삭제된 매체에서 음성파일을 복원하여 그 내용을 확인할 수 있도록 하는 과정이 필요하다.

일반적인 파일 복원은 파일 카빙(file carving)<sup>[1]</sup>이라는 기법이 적용돼 수행되어 왔다. 파일 카빙은 파일 시스템 상에서 삭제된 파일을 복원하는 기법이다. 하지만 파일 카빙 기법을 적용하더라도 음성파일을 온전히 복원하지 못해 재생이 어려운 경우가 발생할 수 있다. 예를 들어, 파일 시스템에서 음성파일이 삭제된 뒤, 덮어쓰기가 발생하면 해당 구간의

<sup>†</sup>Corresponding author: Ha-jin Yu(hjyu@uos.ac.kr)  
School of Computer Science College of Engineering, University of Seoul, 163 Siripdae-ro, Dongdaemun-gu, Seoul 02504, Republic of Korea  
(Tel: 82-2-6490-2448, Fax: 82-2-6490-2444)

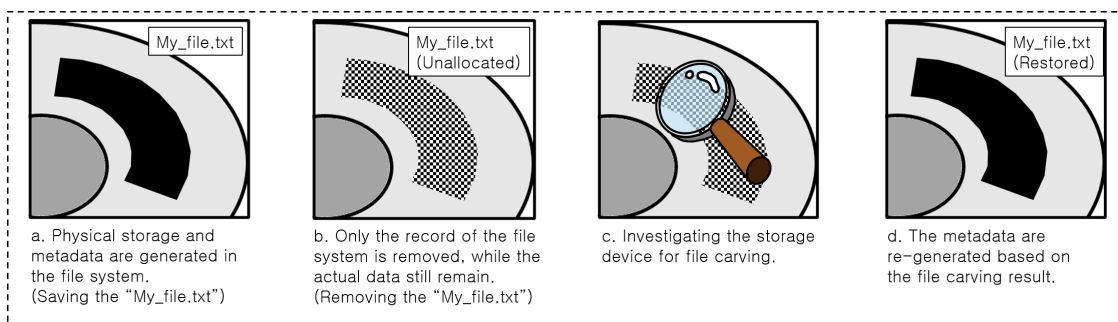


Fig.1. The flow of file restoration based on file carving.

정보가 손실돼 온전한 음성파일이 복원되지 않을 수 있다. 또한, 복원되지 못한 구간이 음성파일 재생에 필수적인 부분(웨이브 파일의 헤더 등)이라면 손상된 부분에 의해 음성파일이 재생되지 않을 수 있다. 이와 같이 손상된 음성파일을 복원하기 위해서는 기존의 파일 복원과 다른 개념의 새로운 복원 기법이 도입되어야 한다. 새로운 복원 기법은 파일의 손상되지 않은 부분을 토대로 손상된 부분의 정보를 유추하는 것이다. 이와 같은 복원 과정을 거치게 되면 기존의 복원 기법(파일 카빙)으로는 복원하지 못했던 손실된 정보까지도 복원이 가능할 것이다. 본 논문에서는 새로운 복원 기법을 수행하는 과정에서 필요한 작업이지만 사람이 직접 수행하기에는 무리가 있는 작업을 심층 신경망(deep neural network)을 적용해 자동화시키는 연구를 진행하였다.

본 논문의 다음 장에서는 기존의 복원 기법인 파일 카빙에 대해서 설명하고 3장에서 심층 신경망의 활용 방안을 제안한다. 제안한 심층 신경망 기반의 활용 방안의 성능을 확인하는 실험을 4장과 5장에서 보이고 6장에서 결론을 맺는다.

## II. 파일 카빙

기존의 파일 복원 기술들은 대부분 파일 카빙 기반의 기술들을 의미한다. 파일 카빙은 파일 시스템에서 삭제된 파일을 구조와 내용물을 기반으로 복원시키는 기술이다.<sup>[1]</sup> 파일 시스템은 사용자가 사용하는 모든 파일들을 관리하는 시스템이다. 사용자가 특정 파일을 저장하면, 파일 시스템은 파일이 물리적으로 저장된 위치, 생성한 시각 등의 정보가 담긴

메타데이터를 생성해 저장하고 파일이 저장된 위치에 다른 파일을 덮어씌울 수 없도록 한다[Fig. 1(a)]. 그 뒤, 사용자가 파일을 삭제하면, 파일 시스템은 앞서 생성한 메타데이터를 삭제하고 해당 파일이 저장된 위치에 다른 파일을 덮어씌울 수 있도록 설정을 변경한다[Fig. 1(b)]. 여기서 파일 카빙은 사용자가 파일을 삭제해 메타데이터가 삭제되었더라도, 메타데이터 없이 실제 파일이 저장된 위치를 탐색해 삭제된 파일을 읽어 들이는 것을 의미한다[Fig. 1(c)]. 따라서 파일 카빙은 손실된 정보를 복원하지 않고 파일 시스템 상에서 삭제돼 접근이 어렵던 파일을 다시 접근 가능하도록 복원하는 기술이라고 해석할 수 있다. Fig. 1은 앞서 설명한 파일 카빙 기반의 파일 복원 과정 보여준다.

하지만 이와 같은 복원 방식으로는 파일이 삭제된 뒤에 덮어쓰기가 발생하지 않은 경우에만 온전하게 파일을 복원할 수 있다. 예를 들어 음성파일의 한 종류인 웨이브 파일이 삭제된 뒤 덮어쓰기가 발생해 음성 재생에 꼭 필요한 헤더 정보가 손실된 경우 파일 카빙 기법으로 웨이브 파일을 복원하더라도 정상적으로 음성을 재생하기 어렵다.

## III. 심층 신경망의 활용

헤더가 손상돼 재생이 어려운 웨이브 파일은 헤더 이외의 부분(부호화된 신호)을 활용해 손상된 정보를 유추하는 새로운 개념의 복원 기법이 적용되어야 한다. 새로운 복원 과정은 기존의 파일 카빙을 통해 수행되었던 복원 과정과는 다르며, 파일 카빙을 통해서 복원하지 못했던 부분까지 복원할 수 있는 방

법이다. 새로운 복원 과정에 필요한 작업들은 사람이 직접 수행하기에는 너무 많은 시간과 비용이 소요된다. 본 논문에서는 심층 신경망을 활용하여 이러한 작업들을 자동화하고, 새로운 복원 기법을 개발하는 두 가지 방안을 제안하였다.

### 3.1 음성, 비음성 분류를 통한 헤더 정보 유추

헤더가 손상된 음성파일은 다음과 같은 과정을 통해 복원될 수 있다. 먼저 손상된 음성파일에서 기존에 존재할 수 있는 헤더 정보를 생성한다. 예를 들어, 음성파일의 헤더에 양자화 지수 정보만이 저장된다면, 음성파일에 사용될 수 있는 모든 양자화 지수에 해당하는 헤더 정보를 생성해야 한다. 그 뒤, 생성한 헤더 정보를 기반으로 음성파일을 복호화하고 복호화된 신호의 음성, 비음성 여부를 판별한다. 여기서의 음성 신호는 손상되기 이전의 헤더 정보가 제대로 복원돼 정상적인 음성을 확인할 수 있는 신호를 의미하며 비음성 신호는 임의로 생성한 헤더 정보가 손상되기 이전의 헤더 정보와 차이가 있어 잘못 복호화된 신호를 의미한다. 음성 신호를 부호화할 때 사용했던 헤더 정보가 아니면 정상적인 음성 신호가 복호화될 수 없기 때문에, 정상적인 음성 신호가 복호화될 때까지 헤더 정보를 임의로 생성하고 이를 사용해 음성 신호를 복호화하는 과정을 반복한다.

위의 과정을 통해 정상적인 음성 신호인 것으로 확인할 수 있는 경우, 실제 손실된 헤더 정보를 유추하였다고 할 수 있지만 이를 사람이 직접 수행하게 되면 손상된 음성파일에서 발생할 수 있는 다양한 경우의 수를 고려하여 모든 경우의 수를 대상으로 신호를 복호화하고 청취하여 음성, 비음성 여부를 확인하는 작업을 반복해 수행하여야 한다. 하지만 실제 음성파일의 부호화 방식을 고려하면 발생할 수 있는 경우의 수가 너무 많기 때문에 현실적인 시간 안에 음성파일을 복원하는 것이 불가능할 수 있다. 따라서 본 연구에서는 위의 과정에 심층 신경망을 적용해 음성, 비음성 여부를 자동으로 판별할 수 있도록 시스템을 구성하였다.

실제 사람의 음성이 포함된 음성 신호의 경우에는 주파수 대역에서 포먼트와 같은 특성이 나타나고 비

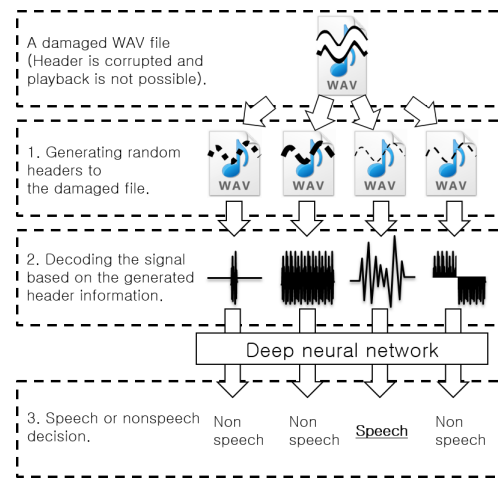


Fig. 2. The flow of speech or nonspeech decision.

음성의 경우, 백색 잡음과 유사한 특성이 나타날 것이라고 가정하고 심층 신경망을 통한 음성, 비음성 분류가 가능할 것이라고 기대하였다. 심층 신경망은 가장 단순한 구조의 앞먹임 네트워크로 구성하였으며 구체적인 동작은 다음과 같다. 먼저 복호화된 신호를 작은 프레임 단위로 분할한 뒤, 각 프레임 별로 주파수 대역의 특성을 잘 나타낼 수 있는 특징을 추출하였다. 추출한 특징을 심층 신경망에 입력하기 위하여 다수의 프레임에서 추출한 특징을 연결하여 사용하였다. 심층 신경망은 입력된 특징에 대하여 이진 분류를 수행하여 입력된 신호가 음성 신호에 가까운지 비음성 신호에 가까운지 분류하게 된다. 각 프레임 단위의 식별 결과를 발생 단위로 취합하여 최종적으로 복호화된 신호가 음성인지 비음성인지를 판별하게 된다. 이와 같은 구조는 심층 신경망을 사용한 음성인식 시스템<sup>[2]</sup>을 참고해 구성하였으며 전체적인 동작은 Fig. 2와 같다.

### 3.2 Long short term memory 기반

#### 음성파일 식별

음성 신호는 음파를 수치화시킨 데이터이다. 따라서 음성 신호에 포함돼 있는 각 샘플들은 독립적으로 존재하지 않고 음파가 가지는 진폭이나 진동수와 같은 특성에 의해 그 값이 결정되며 각 샘플 간의 종속성이 생기게 된다. 음성 신호의 각 샘플들이 가지는 종속성은 다음과 같이 두 가지로 나뉘어서 생각할

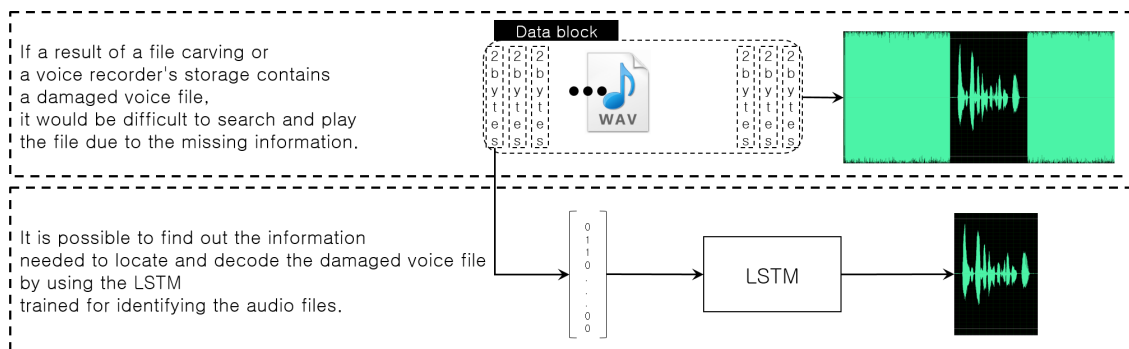


Fig. 3. An example of file restoration using the LSTM.

수 있다. 먼저 짧은 범위 내의 있는 샘플들을 고려하면, 샘플 간의 변화폭이 어느 정도 제한돼 있으며 앞 뒤 샘플들을 참고해 특정 샘플을 예측하는 것도 가능하기 때문에 짧은 간격 종속성이 있다고 할 수 있다. 또한, 샘플 값이 긴 기간 동안 꾸준히 감소하거나 증가하지 않고 증가와 감소를 주기적으로 반복하기 때문에 긴 간격종속성이 있다고 할 수 있다. 이와 같은 특성을 가지고 있는 음성 신호를 부호화해 저장하는 경우, 부호화된 데이터도 음성 신호의 영향으로 데이터 간 종속성이나 특성을 보일 수 있다. 예를 들어, 음성 신호를 16 bit 방식으로 저장한 웨이브 파일에서는 2 bite 단위로 앞서 설명한 종속성이나 나타나며, 8 bit 방식으로 저장한 웨이브 파일에서는 1 bite 단위로 종속성이 나타날 것이다.

따라서 LSTM(Long Short Term Memory)<sup>[3]</sup>와 같이 짧은 간격 종속성과 긴 간격 종속성을 동시에 학습할 수 있는 심층 신경망을 활용하면 음성파일을 모델링할 수 있을 것이다. LSTM을 활용한 음성파일 모델링은 앞서 설명한 음성, 비음성 분류 실험과 다르게 음성 신호를 부호화하는 과정 없이 다음과 같이 진행된다. 먼저 부호화된 음성파일을 일정한 단위로 읽어 들인 뒤, 이를 이진 벡터로 표현하게 된다. 구체적으로 2 bite 단위로 음성파일을 읽어 들이면 16개의 이진값으로 구성된 벡터가 생성된다. 음성파일 별로 생성한 이진 벡터를 LSTM에 입력해 LSTM이 음성파일의 부호화 방식을 식별할 수 있도록 학습시킨다. 이후에 학습시킨 LSTM을 활용하면 음성파일에서 신호를 부호화하는 과정 없이 각 음성파일들을 식별하는 것이 가능하다.

앞에서 설명한 음성파일 식별 LSTM은 다음과 같

이 음성파일 복원에 활용할 수 있다. 먼저 파일 카빙의 결과물로 생성되었거나 음성 녹음기의 저장매체에서 추출한 데이터 블록에 손상된 음성파일이 포함되어 있다고 가정한다. 데이터 블록에 포함된 음성파일을 복원하는 것이 목적이지만 헤더 정보가 손상돼 있기 때문에 데이터 블록에서 음성파일의 위치를 탐색하는 것이 어려울 것이다. 하지만 음성파일 식별 LSTM을 활용하면 데이터 블록을 일정 구간으로 분할하여 각 구간에서 미리 학습한 음성파일 특성이 나타나는지 확인해 음성파일의 위치와 부호화 방식을 동시에 알아내는 것이 가능하다. Fig. 3은 앞서 설명한 LSTM의 활용 방안을 보여준다.

#### IV. 실험 설계

본 논문에서는 제안한 심층 신경망의 활용 방안의 실현 가능성과 효율 등을 확인하기 위해 다음과 같이 실험을 설계해 진행 하였다. 본 논문에서 다루고 있는 실험은 기존의 파일 카빙 기반의 복원이 불가능한 상황을 가정하여 설계한 실험이기 때문에 기존의 복원 시스템은 고려하지 않고 실험을 진행하였다. 실험에 사용한 모든 심층 신경망은 Theano<sup>[4,5]</sup> 환경에서 구현하였다.

##### 4.1 음성 데이터베이스

음성, 비음성 분류와 음성파일 식별 실험을 위해 동일한 음성 데이터베이스를 사용하였으며, 실험의 목적에 맞도록 데이터를 변환하였다. 한국전자통신연구원에서 배포한 한국어 중가마이크 화자인식용

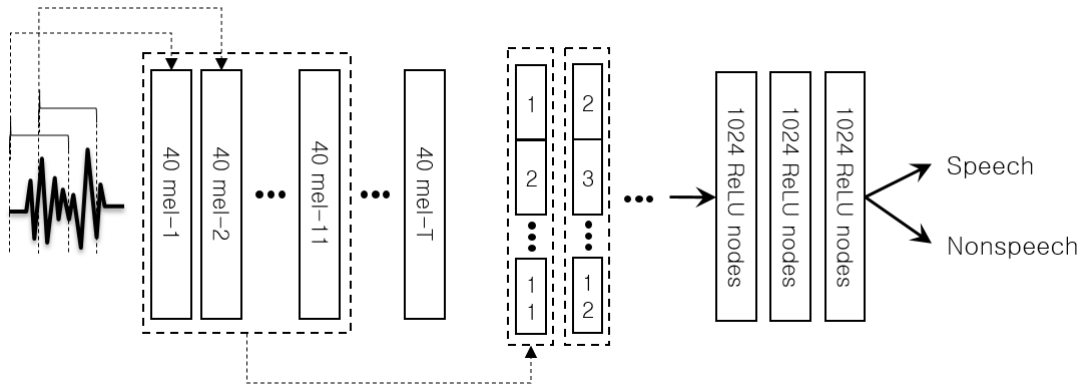


Fig. 4. The flow of speech or nonspeech decision.

음성 데이터베이스를 학습용, 5~6세 아동을 위한 놀이동산 따라 말하기 선별 검사 데이터베이스를 평가용으로 사용해 실험을 진행하였다. 두 종류의 음성 데이터베이스를 사용해 데이터베이스를 구성하고 있는 화자의 연령대, 채널, 음소 정보 등을 완전히 분리시키고자 하였다.

#### 4.2 음성, 비음성 분류 실험 설계

음성, 비음성 분류 실험은 웨이브 파일을 기준으로 잘못 복호화된 비음성 신호와 정상적으로 복호화된 음성 신호를 식별하는 것이 가능한지 확인하는 실험이다. 여기서 비음성 신호는 3.1 장에서 설명한 헤더 정보를 유추하는 과정에서 잘못된 헤더 정보를 웨이브 파일에 대입하는 경우를 가정한 것이다.

음성 신호 학습에 16 kHz, 16 bit, 모노 형식의 웨이브 파일을 정상적으로 복호화된 신호를 사용하였고 비음성 신호 학습에는 앞의 웨이브 파일을 8 bit 무부호 방식으로 잘못 복호화된 신호를 사용하였다. 음성 신호 평가에는 16 kHz, 16 bit, 모노 형식의 웨이브 파일을 정상적으로 복호화된 신호와 음성 신호에 백색 잡음을 10 dB과 0 dB SNR(Signal Noise Ratio, 신호 대잡음비)으로 삽입한 신호를 사용하였다. 잡음이 없는 환경에서 수집한 음성 신호만을 학습한 심층 신경망이 백색 잡음이 포함된 음성 신호에 대해 얼마나 강인한 인식 성능을 보이는지 확인하기 위해, 평가에서만 백색 잡음을 삽입한 음성 신호를 사용하였다. 비음성 신호 평가에는 앞의 웨이브 파일을 8 bit 무부호 방식, 16 bit big endian 방식, 8 bit  $\mu$ raw 방식, 8 bit A raw 방식으로 잘못 복호화된 신호를 사용하였

다. 8 bit 무부호 방식으로 잘못 복호화된 비음성 신호만으로 학습한 심층 신경망이 다른 종류의 비음성 신호에 대해 어느 정도의 일반화 성능을 보이는지 확인하기 위해 다양한 방식으로 비음성 신호를 생성해 사용하였다. 웨이브 파일에 정상적인 헤더 정보를 생성한 한 가지 경우에서만 음성 신호를 확인할 수 있기 때문에 한 종류의 음성 신호 부호화 방법만을 가정하였다.

미리 생성한 음성, 비음성 신호를 심층 신경망 분류기에 학습시킨 뒤에, 5 s 단위의 신호를 심층 신경망에 입력해 음성, 비음성 여부를 판별하였다. 심층 신경망에 입력되는 신호에 25 ms 윈도우를 10 ms씩 이동시켜가며 40차 mel-filterbank 특징을 추출하였다. 11개 윈도우에서 추출한 특징을 연결하여 심층 신경망에 입력할 440차 특징을 생성하였다. 심층 신경망은 입력층과 세 개의 은닉층, 출력층으로 구성된 앞먹임 네트워크 형식으로 구성하였다. 세 개의 은닉층은 각각 1024개의 노드를 포함하고 있으며 각 노드는 ReLU(Rectified Linear Unit)<sup>[6]</sup> 함수에 의해 활성화된다. 출력층의 두 노드는 각각 음성 신호와 비음성 신호를 의미하며 softmax 함수에 의해 활성화된다. 출력층에서 계산한 NLL(Negative Log Likelihood) 값을 학습 데이터에 대해 최소화시키도록 심층 신경망을 반복 학습시켰다. 학습율은 0.1로 하고 drop-out 비율을 0.5로 적용하여 학습을 수행하였다. 음성, 비음성 분류 실험의 동작은 Fig. 4와 같다.

#### 4.3 음성파일 부호화 방식 식별 실험 설계

음성파일 부호화 방식 식별 실험을 위해 학습용

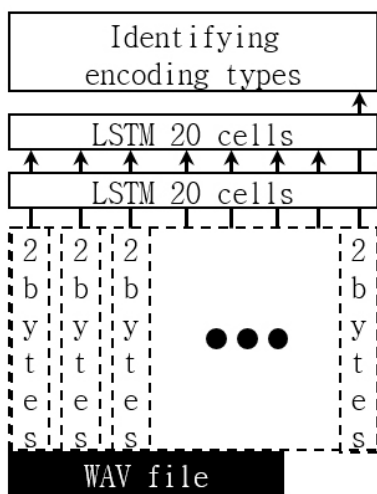


Fig. 5. The flow of audio file identification.

데이터베이스와 평가용 데이터베이스에서 16 bit big-endian, 16 bit little-endian, 8 bit mu-law, 8 bit A-law 방식으로 부호화된 웨이브 파일을 생성하였다. 학습용 데이터베이스에서 생성한 웨이브 파일을 사용해 네 종류의 웨이브 파일 특성을 LSTM에 학습시킨 뒤, 평가 데이터베이스의 네 종류 파일을 식별하여 식별 정확도를 측정하였다. 실험에 사용한 LSTM은 16개의 노드를 포함하는 입력층, 20개의 cell을 포함하는 2개의 은닉층, 4개의 노드를 포함하는 출력층으로 구성하였다. 학습율은 0.01로 하고 drop-out 비율을 0.5로 적용하여 학습을 수행하였다. 음성파일 분류 실험의 동작은 Fig.5와 같다. 본 실험은 Fig.3에서 설명하고 있는 LSTM을 활용한 파일 복원이 가능한지 확인하는 실험으로서 실제 복원과 차이가 있을 수 있다. 실제 복원 과정을 가정하여 실험을 진행하는 경우에는 파일의 손상 유형나 크기, 데이터 블록에서의 위치 등 다양한 요소들을 고려해야하기 때문에 본 실험을 통해 간략히 그 가능성만을 확인하고자 하였다.

## V. 실험 결과

### 5.1 음성, 비음성 분류 실험 결과

심층 신경망을 활용한 음성, 비음성 분류 실험의 결과는 Table 1과 같다. Table 1은 헤더 정보가 손실된

Table. 1. The results of speech or nonspeech decision.

	Data type	Accuracy(%)
Matched	16 kHz, 16 bit, mono	100.0
	+ 10 dB white noise	88.3
	+ 0 dB white noise	41.7
Un-matched	+8 bit unsigned decoding	100.0
	+16 bit big-endian decoding	100.0
	+8 bit $\mu$ law decoding	100.0
	+8 bit A-law decoding	100.0

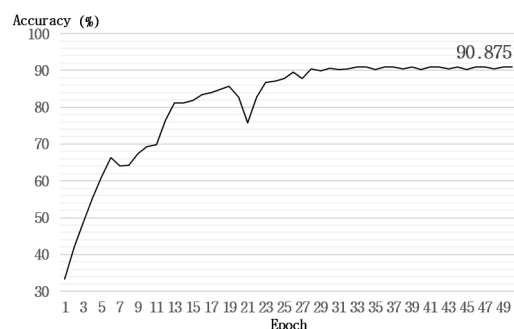


Fig. 6. Identification accuracies for each epoch.

웨이브 파일을 가정하였을 때, 임의로 생성한 헤더 정보가 기존의 헤더 정보와 일치하여 정상적인 음성 신호가 복호화된 경우(Matched)와 헤더 정보가 일치하지 않아 비음성 신호가 복호화된 경우(Un-matched), 각각의 식별 정확도를 보여준다. 실험 결과를 보면 잡음이 삽입되지 않은 음성 신호와 비음성 신호를 100%의 정확도로 분류하는 것을 확인할 수 있다. 다만 백색 잡음이 삽입된 음성 신호에 대한 분류 정확도가 잡음의 크기에 비례해 급격히 감소하는 것을 확인하였다. 이는 잡음이 전혀 삽입되지 않은 음성 신호만을 사용해 심층 신경망을 학습시켰기 때문인 것으로 보인다.

### 5.2 음성파일 부호화 방식 식별 실험 결과

Fig. 6을 보면 LSTM을 반복 학습시키며 진행한 웨이브 파일 식별 실험의 결과를 확인할 수 있다. 실험 결과를 보면 50회 이상 LSTM을 학습시키는 경우, 네 종류의 웨이브 파일을 90% 이상의 정확도로 식별하는 것을 알 수 있다.

## VI. 결론

기존 파일 카빙 기반의 복원 기법은 손실된 정보를 유추할 수 없기 때문에 더욱 효과적인 파일 복원을 위해 손상된 정보까지 유추할 수 있는 새로운 복원 기법에 대한 연구가 필요하다. 본 논문에서는 차후 파일 복원 기법 개발에 도움이 될 수 있는 심층 신경망 활용 방안을 연구하였다. 음성파일에서 손실된 정보를 유추하기 위해 필요하지만 사람이 직접 수행할 수 없는 작업을 심층 신경망이 대신 수행할 수 있는지 확인하는 연구를 진행하였다.

본 논문에서는 음성파일 복원에 활용할 수 있는 두 가지 심층 신경망 활용 방안을 제안하였다. 첫 번째 활용 방안은 심층 신경망으로 음성, 비음성 분류를 자동으로 수행해 음성파일의 헤더 정보를 유추하여 손상된 음성 파일을 복원할 수 있도록 하는 것이다. 두 번째 활용 방안은 음성파일을 복호화하지 않은 상태로 식별할 수 있는 심층 신경망을 학습해 데이터블록에서 손상된 음성파일을 탐색할 수 있도록 하는 것이다. 제안한 두 활용 방안의 실현 가능성과 효율을 확인하는 실험을 웨이브 파일을 대상으로 설계해 진행하였다. 실험 결과, 심층 신경망이 기대한 것과 같이 동작하는 것을 확인해 차후에 완벽한 음성파일 복원 기법을 개발할 때, 활용할 수 있을 것으로 기대할 수 있다.

차후에는 제안한 심층 신경망 활용안을 적용해 파일 복원 기법을 개발하는 연구를 진행할 계획이다. 또한, 실험 결과로 확인할 수 있는 것과 같이 심층 신경망이 음성, 비음성 분류 시 백색 잡음에 취약한 약점을 보완할 수 있는 연구도 함께 진행할 계획이다.

## 감사의 글

이 논문은 2016년 대검찰청 연구용역의 지원으로 수행되었음(과제명: 손상된 음성파일 복원 기법 연구).

## References

1. P. Rainer, S. Tjoa, and P. Tavalato, "Advanced file carving approaches for multimedia files," *JoWUA* **2**,

42-58, (2011).

2. G. Alex, A. Mohamed, and G. E. Hinton, "Speech recognition with deep recurrent neural networks," *ICASSP*, 6645-6649, (2013).
3. H. Sepp and J. Schmidhuber, "Long short-term memory," *Neural computation*, **9**, 1735-1780, (1997).
4. *Theano: new features and speed improvements*, <https://arxiv.org/abs/1211.5590>, 2012.
5. J. Bergstra, O. Breuleux, F. Bastien, P. Lamblin, R. Pascanu, G. Desjardins, J. Turian, D. Warde-Farley, and Y. Bengio, "Theano: a cpu and gpu math expression compiler," in *Proc. Python for Scientific Computing Conference*, 3-10, (2010).
6. V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," *ICML*, 807-814, (2010).

## 저자 약력

### ▶ 허 희 수 (Hee-Soo Heo)



2013년 2월: 서울시립대학교 컴퓨터과학부 학사

2013년 ~ 현재: 서울시립대학교 컴퓨터과학부 석박사통합과정

### ▶ 소 병 민 (Byung-Min So)



2010년 2월: 서울시립대학교 컴퓨터과학부 학사

2012년 8월: 서울시립대학교 컴퓨터과학부 석사

2012년 ~ 현재: 대검찰청 수사관

### ▶ 양 일 호 (IL-Ho Yang)



2008년 2월: 서울시립대학교 컴퓨터과학부 학사

2010년 2월: 서울시립대학교 컴퓨터과학부 석사

2010년 ~ 현재: 서울시립대학교 컴퓨터과학부 박사과정

▶ 윤 성 현 (Sung-Hyun Yoon)



2015년 8월: 서울시립대학교 컴퓨터과학부  
학사  
2015년 ~ 현재: 서울시립대학교 컴퓨터  
과학부 석사과정

▶ 유 하 진 (Ha-Jin Yu)



1990년 2월: KAIST 전산학과 학사  
1992년 2월: KAIST 전산학과 석사  
1997년 2월: KAIST 전산학과 박사  
1997년 ~ 2000년: LG 전자 전자기술원  
선임연구원  
2000년 ~ 2002년: SL2(주) 연구소장  
2002년 ~ 현재: 서울시립대학교 컴퓨터  
과학부 교수