

순환신경망 기법을 이용한 스파 플랫폼의 시계열데이터 필터링에 관한 연구

유승열¹ · 이재철² · 이종현³ · 황호진⁴ · 이순섭[†]

(Received December 27, 2018 ; Revised January 17, 2019 ; Accepted January 21, 2019)

A study on time series data filtering of spar platform using recurrent neural network

Seung-Yeol Yoo¹ · Jae-Chul Lee² · Jong-Hyun Lee³ · Ho-Jin Hwang⁴ · Soon-Sup Lee[†]

요약: 스마트 선박 (Smart ship)의 개발과 해양 플랫폼의 예지보전 시스템 및 자산 관리 시스템 개발을 위해 방대한 양의 계측 데이터를 실시간으로 분석할 수 있는 기술에 대한 관심이 높아지고 있다. 이러한 계측 데이터를 실시간으로 분석하기 위해서는 계측 데이터의 노이즈를 제거하고 필요한 정보를 추출하여 분석에 용이한 형태로 데이터를 가공하는 과정인 데이터 필터링이 반드시 선행되어야 한다. 기존의 조선 해양 산업에서는 일정기간 이상 데이터를 저장한 후 이에 대한 분석을 실시하여 스펙트럼 기반의 필터링 기법을 많이 이용하였다. 이러한 방법은 실시간 데이터를 분석해야 하는 현 상황에는 적합하지 않아 실시간 데이터를 필터링하기 위한 새로운 기법이 필요한 실정이다. 본 논문에서는 시계열 데이터를 학습하기 위한 딥 러닝 모델인 순환신경망 알고리즘을 이용하여 실시간으로 전송되는 데이터를 필터링하고자 하였다. 실시간으로 계측되는 스파 플랫폼의 계류 장력 값을 필터링하기 위해 순환신경망 알고리즘을 이용한 필터링 모델을 설계하고 그 결과값을 확인하여 실시간 필터링 가능 여부를 확인하였다. 최종적으로 실시간으로 전송되는 데이터를 필터링 하기 위해 순환신경망 알고리즘을 사용하는 것이 적합하다는 것을 확인하였다.

주제어: 데이터 필터링, 순환신경망, 시계열 데이터, 실시간, 필터링 기법

Abstract: There is growing interest in the numerous techniques focused on analyzing vast quantities of measurement data in real time for the development of smart ships along with the development of asset integrity management systems for offshore platforms. To analyze the measurement data in real time, data filtering is used to eliminate the noise in the data and then extract the necessary information to perform a comprehensive data analysis. In the traditional shipbuilding and offshore industry, spectrum-based filtering methods are used because the corresponding data is saved for a certain period and subsequently analyzed. These methods are not suited to the present situation in which real-time data is required to be analyzed. Therefore, a new method for data filtering is required. The objective of this study is to filter data in real time using a recurrent neural network algorithm, which is a deep learning model used for learning time series data. In order to filter the measured mooring tension value of the spa platform in real time, a filtering model comprising a recurrent neural network algorithm was designed, and the results of the data filtering process were verified to confirm the possibility of real-time filtering.

Keywords: Data filtering, Recurrent neural network, Time series data, Real-time, Filtering methods

1. 서론

4차 산업혁명의 핵심 기술인 데이터 센싱 기술과 ICT (Information and communications technology)의 활용도가 커짐에 따라 센서로부터 계측된 데이터를 활용하여 장비나

시스템을 실시간으로 제어 및 관리하기 위한 기술이 주목받고 있다. 최근 들어 조선 해양 산업에서도 스마트 선박 (Smart ship)의 개발과 해양 플랫폼의 예지 보전 시스템 및 자산 관리 시스템 개발을 위해 실시간 데이터 분석이 중요

[†] Corresponding Author (ORCID: <http://orcid.org/0000-0002-9698-6884>): Professor, Department of Ocean System Engineering, Gyeongsang National University, Cheondaegukchi-Gil 38, Tongyeong, Gyeongnam, 53064, Korea, E-mail: gnusslee@gnu.ac.kr, Tel: 055-772-9191

1 M.S. Candidate, Department of Ocean System Engineering, Gyeongsang National University, E-mail: holdcury@gmail.com, Tel: 055-772-9197

2 Professor, Department of Naval Architecture and Ocean Engineering, Gyeongsang National University, E-mail: j.c.lee@gnu.ac.kr, Tel: 055-772-9195

3 Professor, Department of Naval Architecture and Ocean Engineering, Gyeongsang National University, E-mail: gnjuhlee@gnu.ac.kr, Tel: 055-772-9194

4 Principal Researcher, Maritime Safety and Environmental Research Division, Korea Research Institute of Ships and Ocean Engineering, E-mail: hjhwang@kriso.re.kr, Tel: 042-866-3645

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

키워드로 자리매김하고 있다. 스마트 선박은 기존의 선박 기술과 ICT가 융합된 것으로써 침체된 조선 해양 산업의 새로운 활로로 주목받고 있는 차세대 선박이다[1]. 이러한 스마트 선박의 핵심은 선박의 상태를 원거리에서 파악하고 제어하는 것이며 이를 위해 1초, 짧게는 0.5초 간격으로 획득되는 계측 데이터를 실시간으로 분석할 수 있어야 한다. 또한 해양 플랫폼의 경우에는 저유가 문제로 인해 새로운 해양 플랫폼을 설치하기보다 기존의 플랫폼의 수명을 연장하여 활용하려는 시도가 늘어나고 있어, 해양 플랫폼의 현재 상태를 실시간으로 진단하여 이상 유무를 파악하고 탑재장비들의 고장을 사전에 예측하여 유지보수하려는 연구가 활발히 진행 중이다[2][3]. 이러한 연구의 목적은 기존의 유지보수 방법에서 벗어나 실시간으로 발생 가능한 고장을 예측하여 해양 플랫폼을 보다 안전하게 운용하는 것이다. 새로운 유지보수 방법은 방대한 양의 데이터를 실시간으로 분석할 수 있는 기술이 필수적이다.

그러나 현실적으로 계측 데이터를 실시간으로 분석하기에는 많은 어려움이 따른다. 가장 큰 문제점은 계측 데이터의 품질이 불완전하다는 점이다. 센서로부터 계측된 데이터는 그 양이 매우 방대할 뿐만 아니라 전송과정에서 발생하는 지연 문제로 인해 정확한 데이터 수집이 불가능하다. 또한 선박이나 해양 플랫폼 내에 존재하는 각각의 장비들이 복잡하게 연결되어 있기 때문에 계측 데이터에는 여러 복합적인 정보가 혼재하여 수많은 노이즈가 발생한다. 이러한 문제점은 실시간으로 계측한 데이터를 분석 데이터로 사용할 수 없도록 만든다. 데이터 분석을 위해서 계측 데이터의 노이즈를 제거하고 필요한 정보를 추출하여 분석에 용이한 형태로 데이터를 가공하는 과정인 데이터 필터링이 반드시 선행되어야 한다. 따라서 정확한 데이터 분석 결과를 얻기 위해 적합한 필터링 기법을 선정하는 것이 무엇보다도 중요하다.

조선 해양 산업에서는 계측 데이터를 분석하기 위해 계측 대상 혹은 데이터의 형태에 따라 적합한 필터링 기법을 선정하여 데이터를 가공하고 있다. T. I. Fossen *et al.*, D. K. Le *et al.* 와 S. Y. Hwang *et al.*은 각각 파도의 세기, 장애물의 위치, 선박의 위치 데이터를 분석하기 위해 칼만 필터를 이용하여 데이터를 가공하였다[4]-[6]. 이들은 계측 데이터에서 나타난 노이즈가 시간의 영향을 받지 않는 화이트 노이즈라는 가정 하에 필터링을 실시하였기 때문에 필터링 결과는 우수했으나 시계열 데이터를 취급하기에는 한계가 있다. D. A. Tobon-Mejia *et al.*, S. K. Kim *et al.*과 S. I. Ann *et al.*은 스펙트럼기반의 분석기법을 이용하여 각각 데이터를 필터링하였다[7]-[9]. 이들은 불필요한 노이즈를 제거하고 분석에 필요한 정보를 정확하게 추출하였으나 일정량 이상의 데이터를 축적한 후 주파수 영역에서 필터링이 진행되고 있어 실시간으로 전송되는 계측 데이터에 적용하기에는 적합하지 않았다. Y. J. Lee *et al.*는 이벤트 로그 정보를 활용하여 선박 엔진 데이터를 필터링하였다[10]. 이러한 방법은

데이터와 시간의 관계를 고려하여 필터링이 진행되기 때문에 시계열 데이터를 취급하기에 적합하다는 장점이 있으나, 고장으로 판단되는 이벤트를 라벨링하여 필터링을 진행하기 때문에 실시간 분석에는 맞지 않았다. Y. W. Song *et al.*과 K. T. Hong *et al.*은 통계적 기법을 이용하여 데이터를 필터링하여 특히 의미 있는 신호 특징을 추출하는데 큰 효과를 보였다[11][12]. 그러나 통계적 기법의 특성상 데이터를 특정 개수 이상 축적해야 한다는 한계를 보였다. 조선 해양 산업에서는 충분한 양의 데이터를 축적한 후 이를 분석에 활용하고 있어 사용되는 필터링 기법 또한 그에 적합한 기법을 활용하고 있다. 반면 데이터의 연속성이나 시간적 특성이 중요시 되는 분야에서는 사용 기법에 있어 조선 해양 분야와 다소 차이를 보였다. 먼저 이미지 프로세싱을 위해 J. W. Lee와 O. S. Byeon은 퍼지 신경망을 이용하여 데이터를 필터링하였다[13][14]. 영상의 연속성을 고려하여 노이즈를 제거하였으며, 특히 O. S. Byeon의 경우 실시간으로 영상의 잡음과 왜곡을 보정하였다. K. Lee *et al.*와 H. J. Seo *et al.*은 음성데이터에 나타나는 노이즈를 실시간으로 제거하여 음성인식 성능을 높이는 연구를 진행하였다[15][16]. 이들은 신경망 알고리즘을 이용하여 특정 주파수 대역의 노이즈를 실시간으로 제거할 수 있다는 것을 보여주었다. G. K. Lee *et al.*와 I. H. Cho Lee *et al.*은 심진도 분석을 위해 기존의 필터링 소요 시간을 단축하고 보다 효과적인 노이즈 제거를 위해 신경망을 이용하여 데이터를 필터링하였다 [17][18]. 칼만필터와 중앙값필터를 복합적으로 이용하는 기존의 방식과 비교하여 필터링에 필요한 소요 시간이 크게 감소되어 실시간으로 데이터를 필터링하기에 적합하였으며 필터링 결과로 발생하는 데이터 왜곡현상이 줄어들었다.

조선 해양 분야의 장비 또는 시스템에서 계측되는 데이터는 특정 주파수 대역을 가지고 있기 때문에 저역통과필터와 같은 스펙트럼기반 분석기법을 많이 이용하고 있다. 그러나 음성 신호와 심진도 신호 또한 특정 주파수 대역의 신호를 추출하기 위해 필터링을 사용하지만 이를 실시간으로 진행하기 위해 신경망 알고리즘을 사용하여 우수한 결과를 획득하는 것을 볼 때, 조선 해양 분야에서 이러한 방법을 적용 시 계측 데이터를 실시간으로 필터링 할 수 있을 것이라 추측하였다.

본 논문에서는 신경망 알고리즘, 특히 순환신경망 알고리즘을 이용하여 미국 멕시코만에 설치된 스파 플랫폼의 시계열 데이터를 실시간으로 필터링하기 위한 연구를 실시하였고 그 결과를 소개한다. 실시간으로 계측되는 스파 플랫폼의 계류 장력 값을 필터링 하기 위해 순환신경망 알고리즘을 이용한 필터링 모델을 설계하고 그 결과값을 확인하여 실시간 필터링 가능 여부를 확인하고자 한다.

2. 데이터 필터링

데이터 필터링이란 데이터에 존재하는 노이즈를 제거하거나 여러 정보가 섞인 데이터 속에서 필요한 정보를 추출

하는 것으로 데이터를 분석에 용이한 형태로 가공하는 행위를 말한다. 데이터를 필터링 하기 위해서 사용되는 기법은 크게 스펙트럼기반 분석 기법, 재귀적 연산 기법, 기계학습 또는 딥러닝을 이용한 기법으로 분류할 수 있다.

먼저 스펙트럼기반 분석 기법은 특정 양 이상의 데이터를 축적한 후 이를 주파수 영역에서 필터링하는 방법이다. 비교적 간단하고 단순한 알고리즘을 이용하여 강력한 필터링 결과값을 얻을 수 있다. 과거부터 많이 사용해오던 방법이기에 때문에 데이터의 특성에 따라 사용되는 기법이 비교적 정형화 되어 있어 결과에 대한 신뢰도가 높다. 그러나 데이터의 시간적 특성을 고려할 수 없으며 무엇보다 축적된 데이터를 사용하기 때문에 실시간으로 필터링을 진행하는 데에는 어려움이 있다. 스펙트럼기반 분석 기법의 대표적인 예로는 필터 주파수 이상의 값을 모두 제거하는 저역통과필터가 있다.

재귀적 연산 기법은 과거에 측정된 데이터와 새롭게 측정된 데이터의 상관관계를 이용하여 데이터를 필터링 하는 방법이다. 재귀적 연산 기법은 기본적으로 두 데이터 사이의 관계가 선형적 관계라는 가정 하에 이루어지는 방법이며, 측정되는 데이터들이 정규분포를 따르거나 그와 유사한 분포를 띄고 있어야 비교적 정확한 값을 도출할 수 있다. 때문에 비선형시스템을 추정하기 위해 여러 변형식이 존재하나 그 결과값의 신뢰도가 낮다. 대표적인 예로는 칼만필터가 있다.

기계학습 및 딥러닝은 학습을 통해 노이즈가 제거된 데이터 값을 예측하거나 정상 값과 노이즈 값을 분류하여 의미 있는 값을 추출할 수 있다. 특히 딥러닝은 여러 개의 히든 레이어를 이용하여 비선형 상태에 대한 추정이 가능하며, 데이터 사이의 상관관계나 특성을 알 수 없어도 학습을 통이 해를 추정할 수 있다. 그러나 상대적으로 다른 필터링 기법에 비해 연구결과가 적어 최적의 기법 선정에 어려움이 있다.

3. 순환신경망 알고리즘

순환신경망 알고리즘은 딥러닝의 주요 모델 중 하나로 순서가 있는 정보를 입력 데이터로 사용하는 신경망 알고리즘이다. 순환신경망 알고리즘은 음성이나 언어와 같은 시퀀스 데이터, 즉 서로 상관관계가 있는 연속된 데이터를 처리하기에 강점이 있는 알고리즘으로 시계열 데이터를 처리하는데 적합하다.

3.1 순환신경망

Figure 1은 가장 기본적인 순환신경망 알고리즘의 모델을 도식화한 것이다. 순환신경망 알고리즘은 다층퍼셉트론(Multi-layer perceptron)과 같은 순방향 신경망과는 다르게 내부노드에서 순환고리를 통해 이전의 정보가 재입력되며 내부 계층을 연결하는 가중치가 모두 동일하다. 즉, t시간의 입력 데이터 x_t 가 주어지면 가중치 U를 곱하여 내부 노

드로 입력되고 t-1초에 주어진 데이터의 결과값에 대해 내부 계층을 연결하는 가중치 W를 곱한 값이 더해져 활성화 함수로 입력되는 구조를 이루고 있다. 순환신경망 알고리즘은 활성화 함수로 tanh (Hyperbolic tangent) 또는 ReLU (Rectified linear unit)를 사용하며 가중치를 학습하기 위해서 편도함수를 이용한 역전파 알고리즘을 이용한다. 순환신경망 알고리즘은 편도함수의 곱으로 가중치를 추정하기 때문에 참조하게 되는 과거의 데이터 수가 많아지게 되면 곱셈의 결과가 0으로 수렴하거나 매우 큰 값으로 발산할 수 있다.

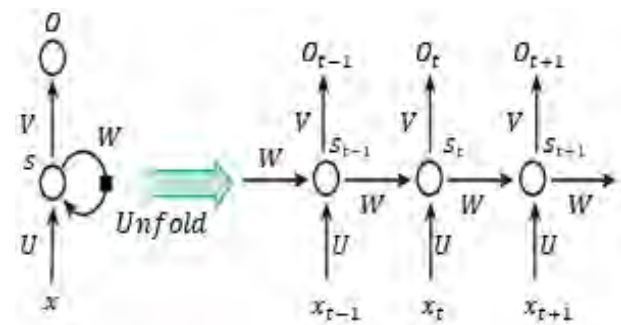


Figure 1: Recurrent neural network

3.2 LSTM (Long Short-Term Memory)

LSTM은 순환신경망 알고리즘의 문제점인 과거 데이터의 수가 많아질 경우 가중치가 0으로 수렴하거나 양의 값으로 발산하는 문제점을 해결하기 위해 수정된 순환신경망 모델로써 기존의 순환신경망 알고리즘 내의 노드를 메모리 셀이라고 불리는 복잡한 구조로 바꾸는 것으로 문제점을 해결하고자 하였다.

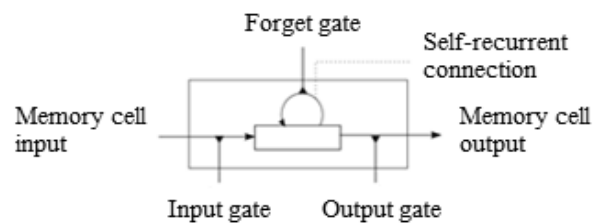


Figure 2: Memory cell of LSTM

Figure 2는 LSTM의 메모리셀을 비교적 단순하게 나타낸 것이다. LSTM의 메모리 셀은 순환입력값을 가지는 내부 기억노드 (Memory cell input)와 세 개의 개폐장치 (Input gate, Forget gate, Output gate)로 이루어진 네 개의 구성요소를 가진다. 망각 게이트 (Forget gate)에서는 시그모이드 (Sigmoid)를 활성화 함수로 가지게 되며 결과값이 0에 가까워 나타날 경우 해당 데이터를 망각 시키는 역할을 한다. 따라서 순환신경망 알고리즘과는 달리 LSTM은 편도함수의 곱이 입력되는 노드 수만큼 계산되는 것을 방지함으로써 보다 정확한 가중치 값을 획득하게 된다.

3.3 순환신경망 알고리즘의 구조

Figure 3에서 보듯이 순환신경망 알고리즘의 구조는 입력 데이터와 결과 값의 상관관계에 따라 결정된다. 가장 간단한 구조는 원 투 원 (One to one)으로 하나의 입력 데이터에 대해 하나의 결과 값을 도출한다.

매니 투 원 (Many to one)은 여러 개의 입력 데이터를 고려하여 하나의 결과 값을 도출하는 구조이다. 본 논문에서는 실시간으로 계측되는 스파 플랫폼의 계류 장력 값을 필터링 하기 위해 순환신경망 알고리즘 모델을 이용하였는데, 120초 주기로 발생하는 조류 (Current)로 인해 발생하는 계류 장력 값을 필터링 하기 위해 120개의 입력 데이터를 이용하여 하나의 결과 값을 도출하였다. 이는 시퀀스 투 벨류 (Sequence to value) 형태이므로 매니 투 원 구조를 사용하였다.

매니 투 매니 (Many to Many)는 여러 개의 입력 데이터를 통해 여러 개의 결과값을 도출하는 것으로 하나의 단어가 입력된 후 그 다음에 오게 될 단어를 예측하는 형식의 언어 인식과 같은 정보처리에 적합한 구조이다.

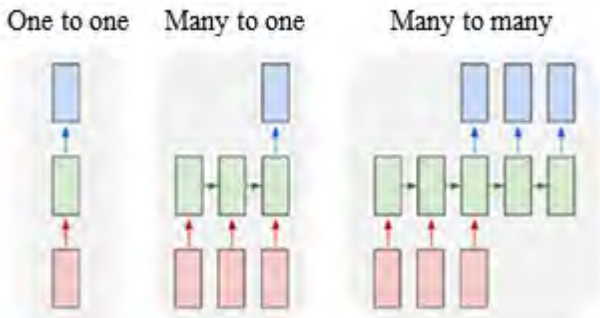


Figure 3: Input/output structure of recurrent neural network

4. 순환신경망 알고리즘의 적용

시계열 데이터에 순환신경망 알고리즘을 적용하여 데이터를 필터링하기 위해 멕시코만에서 운용중인 W사의 G 플랫폼의 계측데이터를 활용하였다. G 플랫폼은 원통형 구조물인 스파 플랫폼이며, 한 지점 당 세 개씩 총 아홉 개의 무어링 라인이 설치되어 플랫폼의 거동을 제어하고 있다. G 플랫폼은 다중도플러식 유속계 (Acoustic doppler current profiler), 에어갭 센서 (Air gap sensor), 위성 항법 보정 시스템 (Differential global positioning system), 자이로컴파스 (Gyro compass), 로드 셀 (Load cells) 등 다양한 계측 센서를 이용하여 해양환경데이터와 G 플랫폼의 거동데이터를 1초 간격으로 24시간, 1년 365일을 쉬지 않고 계측하고 있다. Figure 4는 순환신경망 알고리즘을 이용하여 계측 데이터를 필터링하고 그 성능을 확인하기 위한 과정을 도식화 한 것이다.

4.1 대상 데이터 선정

스파 플랫폼은 파도, 조류, 바람 등의 극심한 환경 속에서 무어링 라인을 이용하여 거동을 제한함으로써 안정적으로 원유 및 천연가스를 생산한다. 만약 스파 플랫폼의 무어링

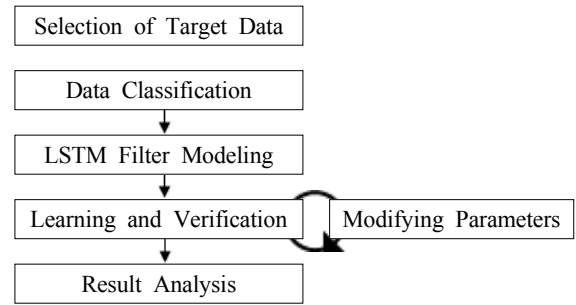


Figure 4: Modeling Process of LSTM filter

라인 (Mooring line)에 문제가 발생할 경우 대형사고로 이어져 막대한 손실이 발생할 수 있으며 최악의 경우 플랫폼이 전복 될 수도 있다. 특히 G 플랫폼이 설치된 멕시코 만의 경우 강력한 루프 커런트 (Loop current)로 인해 원통형 구조물이 조류의 진행방향과 수직된 방향으로 진동하게 되는 VIM (Vortex induced motion) 현상이 발생한다. VIM 현상은 무어링 라인의 피로손상 및 피로파괴를 일으키는 매우 위험한 현상으로 실제로 W 사는 VIM으로 인한 무어링 라인의 피로손상정도를 예측하기 위해 많은 노력을 기울이고 있다.

본 연구에서는 이러한 이유로 인해 G 플랫폼의 여러 계측 데이터 중 무어링 라인의 텐션 데이터를 대상데이터로 선정하였다. 연구 진행을 위해 2014년 10월부터 2018년 2월까지 약 41개월간의 무어링 라인 텐션에 대한 계측데이터인 로우 데이터 (Raw data)와 W사에서 데이터 분석을 위해 사용하고 있는 엔지니어링 데이터 (Engineering data)를 수집하였다. 수집된 데이터 중 2014년부터 2016년까지의 데이터를 학습데이터 및 테스트데이터로 사용하여 필터링 모델을 만들었으며 2017년부터 2018년까지의 데이터를 완성된 필터링 모델에 실시간으로 입력하여 필터링 결과값을 획득하였다. 순환신경망 알고리즘을 이용하여 데이터를 필터링 하기 위해 로우 데이터를 입력 데이터로 사용하였으며 W사의 엔지니어링 데이터와 필터링 결과값을 비교하였다.

4.2 순환신경망 모델링

순환신경망 필터의 역할은 조류의 영향에 의해 발생된 피로손상정도를 예측할 수 있도록 무어링 라인 텐션에 대한 계측 데이터를 실시간으로 필터링하는 것이다. Table 1은 순환신경망 필터 모델의 아키텍처를 정리한 것이다. W사에서는 G 플랫폼에 영향을 미치는 조류의 주기를 120초로 판단하여 분석을 실시하고 있어 매니 투 원 구조를 선택한 후 입력 데이터의 개수를 120개로 선정하였으며 내부 노드 수는 60개로 선정하였다. 단순한 형태의 순환신경망 알고리즘을 사용하기에는 참조해야할 과거 데이터의 수가 너무 많아 기울기 소실의 문제가 발생할 위험이 있어 사용 알고리즘은 개선된 순환신경망인 LSTM을 사용하였다. 활성화 함수는 tanh를 사용하였고 과적합을 방지하기 위해 드롭아웃 (Dropout) 레이어를 사용하였으며 드롭아웃값은 일반적으로 많이 사용하는 값인 30%로 설정하였다. 레이어의

층수가 늘어날수록 보다 깊은 추론이 가능해지기 때문에 오차율을 최대한 줄이고자 단층부터 순차적으로 레이어의 수를 늘려가면서 결과값을 확인하였다. 그러나 레이어 층의 수가 4개가 넘어가는 순간부터 오차율 대비 학습 소요 시간이 급격하게 증가하여 LSTM 레이어의 수는 3개로 선정하였다. 학습 횟수는 5000회로 선정하였고 동일한 값이 10회 이상 도출 시 학습이 조기 종료되도록 설정하였다.

Table 1: Architecture of LSTM filtering model

Hidden layer	Algorithm	Drop out	Input	Output
1st	LSTM	30%	(1,60,1)	(1,60,60)
2nd	LSTM	30%	(1,60,60)	(1,60,60)
3rd	LSTM	30%	(1,60,60)	(1,60)
4th	Dense	-	(1,60)	(1,1)

Figure 5는 필터링 모델을 도식화한 것이다. 입력층 (Input layer)에서 실시간으로 계측된 데이터와 그로부터 120초 전까지의 과거 데이터가 입력된다. 입력된 데이터는 3개의 은닉층 (Hidden layer)을 거쳐 4번째의 은닉층에서 120개 데이터의 상관관계를 하나의 값으로 도출한다. 최종적으로 출력층에서 필터링된 값을 얻는다. 1층부터 3층까지의 은닉층에서 나타나는 점선은 상태유지모드를 나타낸 것으로, 시계열 데이터의 주기성을 고려하기 위해 학습과정에서 상태유지모드 활성화시켜 현재학습의 상태가 다음 학습 시작 시에 초기 상태로 전달되도록 하였다.

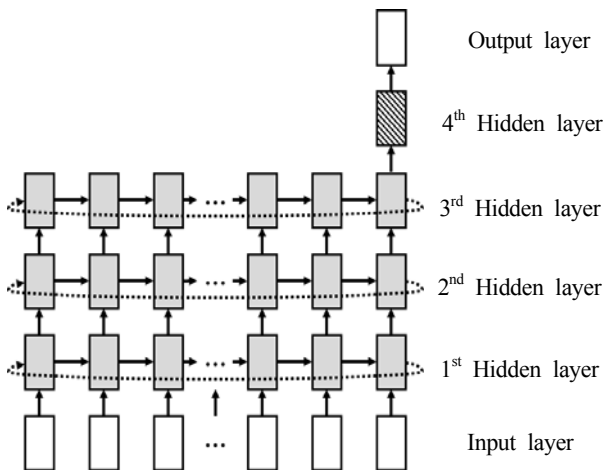
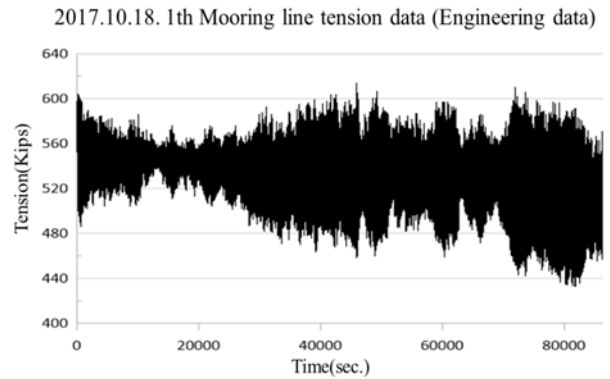


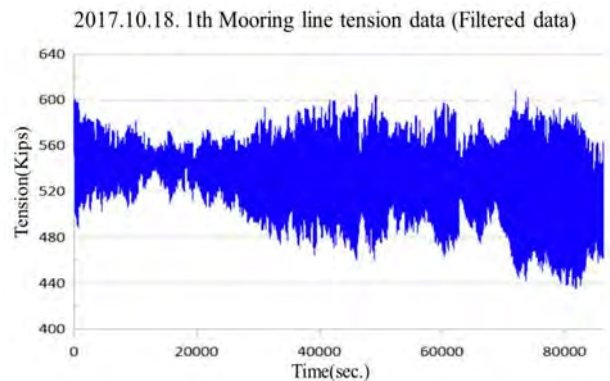
Figure 5: LSTM filtering model

5. 적용 결과

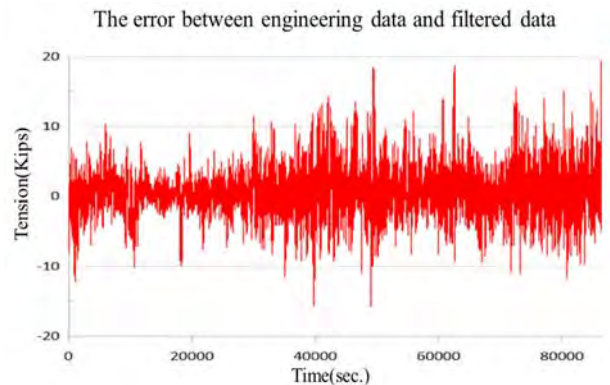
Figure 6는 9개의 무어링 라인 중 1번 무어링 라인의 2017년 10월 18일 하루 동안 계측된 텐션 데이터에 대해 필터링을 진행한 결과이다. 2014년 10월, 2015년 10월, 2016년 10월의 데이터를 학습데이터로 이용하여 순환신경망 모델의 가중치를 결정한 후에 2017년 10월의 데이터를 검증 데이터로 사용하였다.



(a) Engineering data of mooring line tension



(b) Filtered data of mooring line tension



(c) The error between engineering data and filtered data

Figure 6: Comparison of engineering data and filtered data

24시간동안 측정된 텐션데이터에 대한 엔지니어링 데이터는 일정한 패턴을 찾을 수 없으며 특정 구간에서는 증가 또는 감소, 임펄스와 같은 현상이 나타난다.

순환신경망 필터의 실시간 필터링 성능 확인을 위해 2017년 10월 18일 하루, 총 86,400개의 로우 데이터가 1초 간격으로 순환신경망 필터에 입력되도록 하여 데이터를 필터링 하였으며 필터링 결과를 엔지니어링 데이터와 비교하였다. 10월 18일 하루 동안 계측된 데이터 대한 순환신경망 필터의 결과값은 엔지니어링 데이터와 0.47%의 평균 오차율을 보였으며 값의 차이가 가장 많이 나는 시점은 데이터의 끝 지점인 86399초로, 19.295 Kips 만큼의 차이를

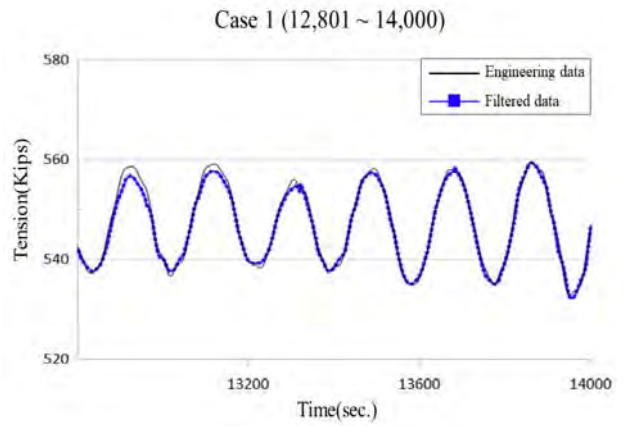
보였다. 평균오차율의 크기로 볼 때, 순환신경망 필터의 결과값을 사용하여 무어링 라인 텐션 데이터를 분석하여도 무리가 없을 것으로 판단되었다. 한편 오차값의 차이가 데이터의 끝 지점에서 가장 크게 나타난 이유는 W사에서 엔지니어링 데이터를 획득하기 위해 사용된 방법에 있는 것으로 판단된다. W사는 엔지니어링 데이터 획득하기 위해 계측 데이터를 하루 간격으로 분류한 후 이를 저역통과 필터로 필터링 하였다. 때문에 2017년 10월 18일 23시 59분 59초에 해당하는 엔지니어링 데이터의 값 (537.3470 Kips) 과 다음 날의 첫 번째 데이터인 2017년 10월 19일 0시 0분 0초의 엔지니어링 데이터의 값(516.2632 Kips)을 확인해보면 21.0838 Kips의 차이를 확인할 수 있다. 엔지니어링 데이터의 변동 폭의 평균값이 0.0065 Kips이고, 최대값이 8.8939 Kips인 것을 볼 때 이러한 차이는 오류를 발생시키는 데 충분하다.

5.1 구간별 분석

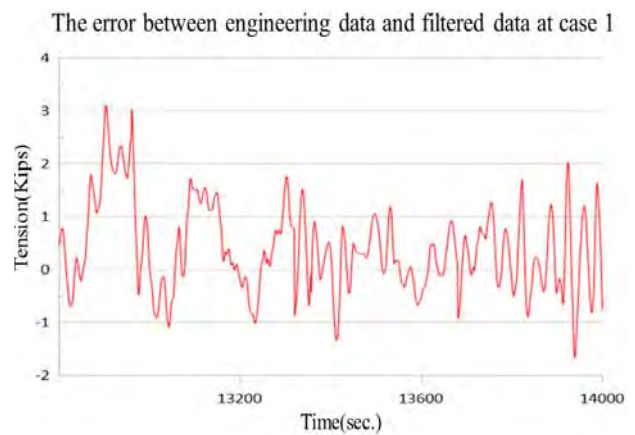
필터링 성능 확인을 위해 특정 구간별로 필터링을 진행하여 결과값을 분석하였다. 구간은 5가지로 분류하였다. 각각의 구간은 가장 작은 진폭을 보이는 구간, 가장 큰 진폭을 보이는 구간, 진폭이 일정하게 유지되는 구간, 증가 구간, 감소 구간으로 나누어 Case 1 부터 Case 5까지 라벨링을 하였다. 각각의 구간 길이는 조류의 주기와 패턴의 형상을 고려하여 1,200초로 결정하였다. 각 Case 별로 학습이 완료된 순환신경망 필터에 1200개의 데이터가 실시간으로 입력되도록 한 후 그 결과 값을 해당 구간의 엔지니어링 데이터와 비교 및 분석하였다.

Case 1 은 가장 작은 진폭을 보이는 구간에 대한 분석이며, Case 2는 가장 큰 진폭을 보이는 구간에 대한 분석이다. Case 3은 진폭의 변화가 크지 않고 비교적 일정하게 유지되는 구간을 분석한 것이고, Case 4 는 진폭이 증가하는 구간에 대해 분석을 실시한 결과이며 Case 5는 진폭이 감소하는 구간에 대한 분석이다.

Figure 7은 Case 1 (12,801 sec. ~ 14,000 sec.)에 대해 순환신경망 필터의 성능을 확인하기 위한 것이다. 1200개의 데이터를 학습이 완료된 순환신경망 필터를 이용하여 필터링한 후 엔지니어링 데이터와 비교 및 분석하였다. Figure 7 (a)는 엔지니어링 데이터와 필터링된 데이터의 형상비교를 나타낸 것이며, Figure 7 (b)는 엔지니어링 데이터에 대한 필터링된 데이터의 오차값을 나타낸 것이다. 평균 오차율은 0.13 % 로 나타났다. 최대 오차와 최소 오차를 구하기 위해 계산된 오차값에 절댓값을 취한 결과 12,904초에서 가장 오차가 크게 나타났으며 그 값은 3.107 Kips 이다. 오차의 최솟값은 13,521초의 0.00085 Kips 로 매우 작게 나타났다. 전체적으로 필터링 결과가 엔지니어링 데이터와 매우 유사하게 나타났으나 진폭의 최고점에서 다소 값이 작게 나타나는 현상을 보였다.



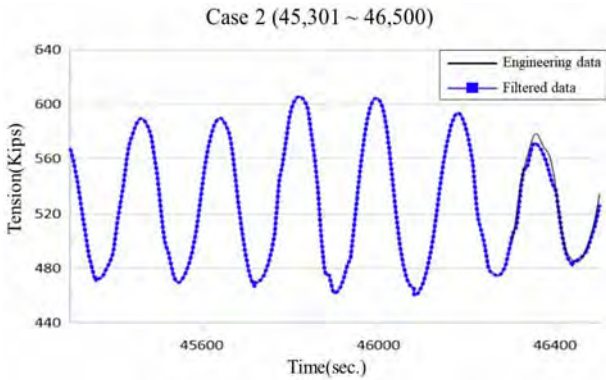
(a) Comparison of the result at case 1



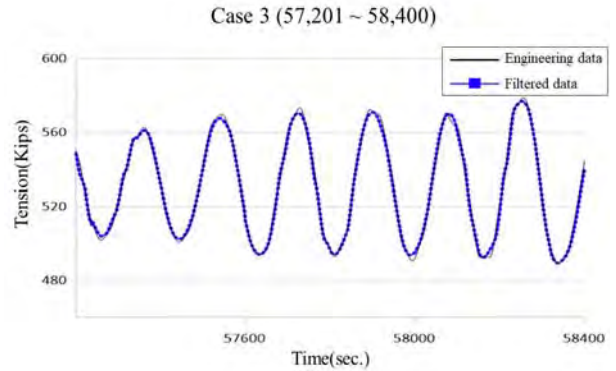
(b) The error between engineering data and filtered data at case 1

Figure 7: Result of case 1

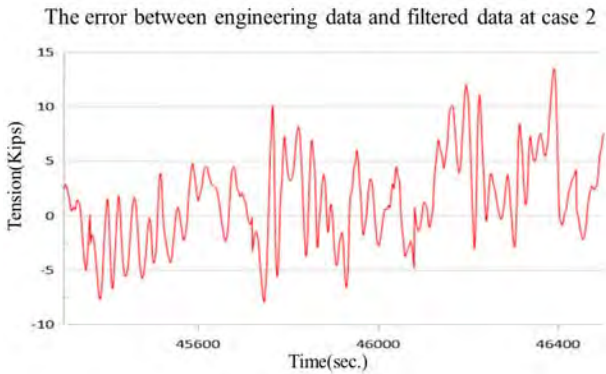
Figure 8은 Case 2 (45,301 sec. ~ 46,500 sec.)에 대한 필터링 결과를 엔지니어링 데이터와 비교한 것이다. Figure 8 (a)는 엔지니어링 데이터와 필터링된 데이터의 형상비교를 나타낸 것이며, Figure 8 (b)는 엔지니어링 데이터에 대한 필터링된 데이터의 오차값을 나타낸 것이다. 평균 오차율은 0.62 % 이며, 46,390초에서 최대 오차인 13.620 Kips 가 나타났다. 최소 오차는 45,650초에서 나타난 0.00813 Kips 이다. Case 1에 비해 상대적으로 오차가 크게 나타났다. 가장 큰 오차가 발생한 46,390초 지점의 형상을 비교해 보면 엔지니어링 데이터는 상대적으로 뾰족한 형상을 하고 있는 것에 비해 필터링 결과는 둥근 형상을 하고 있다. 순환신경망의 알고리즘의 특성상 이전 데이터의 경향성을 반영하기 때문에 특정 지점에서 값이 급격하게 변하는 부분이 매끄럽게 처리된 결과로 인한 오차로 판단된다. Case 2의 그래프를 확대해서 확인하면 진폭이 최대치가 나타난 후 감소하는 시점부터 극값의 형상에 급격한 변형이 생기는데 이로 인해 일정한 패턴으로 돌아오기까지 상대적으로 오차값이 다소 크게 발생하는 것을 확인하였다.



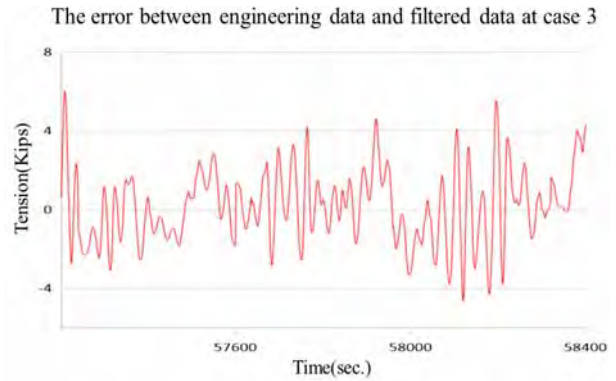
(a) Comparison of the result at case 2



(a) Comparison of the result at case 3



(b) The error between engineering data and filtered data at case 2



(b) The error between engineering data and filtered data at case 3

Figure 8: Result of case 2

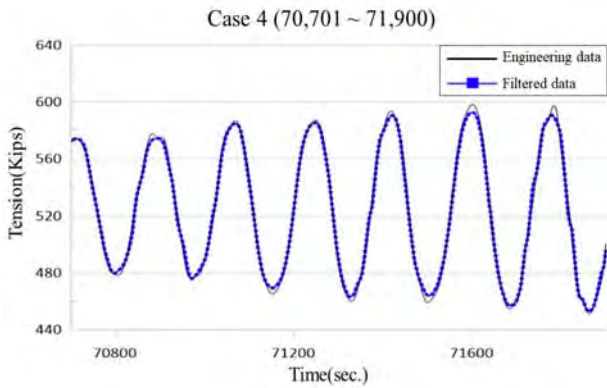
Figure 9: Result of case 3

Figure 9은 Case 3 (57,201 sec. ~ 58,400 sec.)에 대해 순환 신경망 필터의 성능을 확인한 결과를 나타낸 것이다. Figure 9 (a)는 엔지니어링 데이터와 필터링된 데이터의 형상비교를 나타낸 것이며, Figure 9 (b)는 엔지니어링 데이터에 대한 필터링된 데이터의 오차값을 나타낸 것이다. 평균 오차율은 0.28 % 이며, 최대 오차는 57,209초에서 6.620 Kips 로 나타났다. 최소 오차는 57,571초에서 나타난 0.00014 Kips 이다. 진폭의 최고점과 최소점에서 상대적으로 오차가 크게 나타났으며 Case 2와 마찬가지로 엔지니어링 데이터의 값이 상대적으로 급격하게 변하는 형상을 보일수록 차이가 발생하였다. 그러나 전체적으로 일정한 진폭을 유지하고 있는 구간이기 때문에 Case 2에 비해서 상대적으로 오차가 작게 나타난 것으로 판단된다.

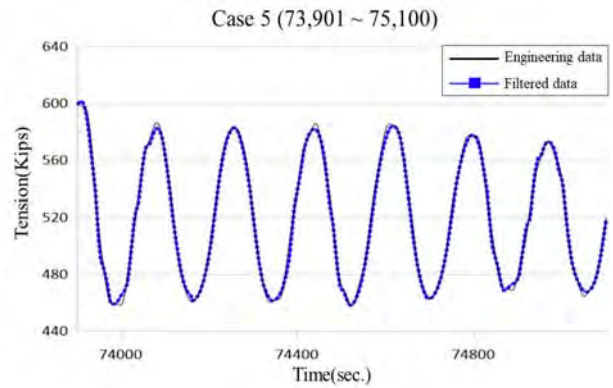
Figure 10은 Case 4 (70,701 sec. ~ 71,900 sec.)에 대해 순환신경망 필터를 사용한 결과와 엔지니어링 데이터를 비교한 것이다. 1200개의 데이터를 학습이 완료된 순환신경망 필터를 이용하여 필터링 한 후 엔지니어링 데이터와 비교 및 분석하였다. Figure 10 (a)는 엔지니어링 데이터와 필터링된 데이터의 형상비교를 나타낸 것이며, Figure 10 (b)는 엔지니어링 데이터에 대한 필터링된 데이터의 오차값을 나타낸 것이다. 평균 오차율은 0.46 % 로 나타났다. 71,736초에서 최대 오차가 나타났으며 그 값은 10.853 Kips 이다. 오차의 최소값은 71,084초의 0.00812 Kips 로 나타났다. 분석

결과 Case 1에 비해 오차값이 크게 나타났으나 Case 4에서 나타나는 엔지니어링 데이터의 최고점 및 최저점의 형상이 대부분 뽀족하게 나타나는 것을 확인하였다. 그래프를 비교해보면 해당 지점을 제외한 구간에서는 상대적으로 오차율이 작게 측정되고 있는 것을 확인 하였다. 다른 증가 구간에 대한 분석을 실시해 본 결과 오차의 크기는 진폭의 증가 상태보다는 엔지니어링 데이터의 형상과 더 큰 연관성을 보이는 것을 확인하였다.

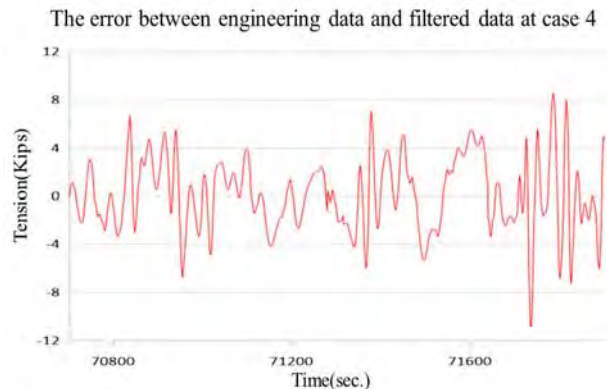
Figure 11은 Case 5 (73,901 sec. ~ 75,100 sec.)에 대해 순환신경망 필터의 필터링 결과를 엔지니어링 데이터와 비교한 것이다. Figure 11 (a)는 엔지니어링 데이터와 필터링된 데이터의 형상비교를 나타낸 것이며, Figure 11 (b)는 엔지니어링 데이터에 대한 필터링된 데이터의 오차값을 나타낸 것이다. 평균 오차율은 0.44 % 이며, 74,382초에서 최대 오차인 11.144 Kips 가 나타났다. 최소 오차는 74,190초에서 나타난 0.00129 Kips 이다. Case 4와 비슷한 결과를 확인할 수 있었다. 가장 큰 오차가 발생한 74,382초 지점의 형상을 보면 엔지니어링 데이터의 값에서 짧은 구간동안 두 개의 극솟값을 확인 할 수 있다. 순환신경망 필터로 필터링된 데이터의 결과를 보면 두 극솟값의 사이를 극솟값으로 하는 곡선이 생성되는 것을 확인 할 수 있다. 동일한 패턴을 보이는 구간에 대해서는 대체로 오차값이 다소 크게 발생하는 것을 확인하였다. 다른 감소 구간에 대한



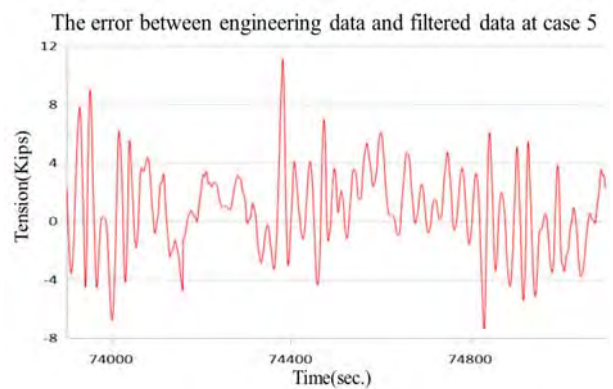
(a) Comparison of the result at case 4



(a) Comparison of the result at case 5



(b) The error between engineering data and filtered data at case 4



(b) The error between engineering data and filtered data at case 5

Figure 10: Result of case 4

Figure 11: Result of case 5 & scale up

분석을 실시해 본 결과 Case 4와 동일하게 오차의 크기는 진폭의 증가 상태보다는 엔지니어링 데이터의 형상과 더 큰 연관성을 보이는 것을 확인하였다. 엔지니어링 데이터의 형상 중 극점 지점에서 급격한 변화를 보여 예측한 형상이 나타나는 지점보다 짧은 시간 동안 두 개의 극점이 잇따라 나타나는 경우에 상대적으로 큰 오차가 발생하는 것을 확인하였다.

5.2 결과 분석

엔지니어링 데이터와 동일하게 하루동안 계측된 데이터를 순환신경망 필터를 이용하여 필터링한 결과 2017년 10월 18일 하루 동안 계측된 데이터에 대한 평균 오차율은 0.47%로 전반적으로 그 값이 유사하게 나타남을 알 수 있었다. 최대 오차는 데이터의 마지막 지점인 86,399초에서 발생하였다. 엔지니어링 데이터를 살펴보면 하루 단위로 로우 데이터를 정형화하기 때문에 하루의 마지막 시간에 해당하는 엔지니어링 데이터와 다음날의 처음 시간에 해당하는 엔지니어링 데이터 사이의 값 차이가 다소 크기 때문에 이러한 결과가 나타난 것으로 판단된다.

실시간으로 필터링을 진행하기 위해 엔지니어링 데이터가 나타내는 그래프의 형상에 따라 구간을 나누어 데이터 필터링을 실시하였다. 그 결과 진폭의 증감보다는 엔지니

어링 데이터의 극점 부근의 형상이 오차 큰 영향을 주는 것을 확인하였다. 이는 과거의 데이터를 이용하여 시간의 흐름에 따른 경향성을 파악하는 순환신경망 알고리즘의 특성 때문인 것으로 판단된다. 엔지니어링 데이터 또한 정확한 값이 아닌 저역통과필터를 이용하여 로우 데이터를 필터링 한 결과이므로, 상대적으로 평탄화가 잘 되어 데이터 형상이 부드럽게 나타나는 순환신경망 필터의 성능이 기존의 엔지니어링 데이터를 획득하기 위한 필터링 기법에 비해 떨어진다고 판단하기는 어려워 보인다.

조선 해양 산업에서 사용하는 해양 플랫폼 무어링 라인의 피로손상정도 예측기법은 텐션의 크기를 일정 간격으로 나누어 라벨링을 한 후 해당 구간에 속하는 텐션이 얼마나 발생했는지 그 횟수를 계산기 때문에 데이터의 오차가 0.65% 미만일 경우는 거의 동일한 결과를 얻게 된다. 본 논문에서 제안한 필터링 모델의 경우 전체구간에서는 0.5%이하의 오차율을 보이고 있으며, 구간별의 최대 오차율은 0.62%로 나타난 것으로 볼 때 엔지니어링 데이터를 이용한 피로손상정도 예측결과와 크게 차이가 나지 않을 것으로 판단되므로 순환신경망 필터를 이용하여 실시간으로 데이터 필터링이 가능하다는 것을 확인할 수 있었다.

6. 결 론

최근 스마트 선박과 해양 플랫폼의 예지 보전 시스템 및 자산 관리 시스템에 대한 관심이 높아짐에 따라 데이터 센싱기술을 이용하여 실시간으로 장비 또는 시스템을 분석하기 위한 기술이 요구되고 있다. 그러나 현실적으로 계측 데이터는 품질이 불완전하여 데이터의 실시간 분석을 위해서 데이터 필터링 또한 실시간으로 이루어져야 한다. 기존의 조선 해양 산업은 정기보전을 통해 장비 및 시스템을 관리하였기 때문에 일정기간 이상 축적된 데이터에 대한 필터링은 많이 이루어졌으나 실시간 데이터 처리에 대해서는 상대적으로 연구가 부족한 실정이다.

본 논문에서는 최근 급속히 발전하고 있는 컴퓨터 하드웨어 기술에 의해 많은 관심이 집중되고 있는 신경망 알고리즘을 통해 데이터를 실시간으로 필터링을 하고자 하였으며 순환신경망 필터링 모델을 설계하여 스파 플랫폼의 계측데이터를 실시간으로 필터링 하였다. 실제 데이터 분석에 사용되는 엔지니어링 데이터와 순환신경망 필터의 필터링 결과를 비교하여 그 성능을 확인하였으며 데이터의 실시간 필터링이 가능하다는 것을 확인하였다.

향후에는 파라미터를 최적화하여 학습시간을 줄이고 동시에 필터링 성능을 높일 예정이며, 거동데이터가 아닌 장비에 대한 계측데이터에 순환신경망 필터를 적용하여 데이터를 실시간으로 필터링 하고자 한다.

후 기

이 논문은 산업통상자원부 ‘산업전문인력역량강화사업’의 재원으로 한국산업기술진흥원(KIAT)의 지원을 받아 수행된 연구(2018년 친환경·스마트 선박 R&D 전문인력 양성 사업, 과제번호: P0001968), 산업통상자원부와 한국산업기술진흥원이 지원하는 경제(광역)협력권산업 육성사업(과제번호: P0004736) 및 2017년도 정부(미래창조과학부)의 지원을 받아 수행된 연구(한국연구재단, 과제번호: NRF - 2017R1C1B5015989) 결과입니다.

Author Contributions

The following statements should be used “Conceptualization, S. Y. Yoo, J. C. Lee, J. H. Lee, H. J. Hwang and S. S. Lee; Methodology, S. Y. Yoo and S. S. Lee; Investigation, J. C. Lee; Resources, J. H. Lee; Modeling, S. Y. Yoo; Verification of Model, S. S. Lee; Data Analysis, S. Y. Yoo and J. C. Lee; Data Interpretation, S. Y. Yoo and J. H. Lee; Verification of Result, H. J. Hwang; Data Curation, S. Y. Yoo, J. C. Lee and J. H. Lee; Writing—Original Draft Preparation, S. Y. Yoo; Writing—Review & Editing, S. Y. Yoo and S. S. Lee; Visualization, J. C. Lee and J. H. Lee; Supervision, S. S. Lee; Project Administration, S. S. Lee; Funding Acquisition, S. S. Lee”.

References

- [1] K. H. Kang, “IoT based equipment of smart ship system,” Bulletin of the Society of Naval Architects of Korea, vol. 55, no. 3, pp 26-28, 2018 (in Korean).
- [2] M. B. Gerner, S. R. Perryman, J. F. Geyer, and J. T. von Aschwege, “Marine monitoring of Gulf of Mexico deepwater floating systems,” Proceedings in Offshore Technology Conference, Houston, Texas, USA, pp. 1-6, 2007.
- [3] P. Elman, J. Bramande, E. Elletson, and K. Pinheiro, “Reducing uncertainty through the use of mooring line monitoring,” Proceedings in Offshore Technology Conference, Rio de Janeiro, Brasil, pp. 1-9, 2013.
- [4] T. I. Fossen and T. Perez, “Kalman filtering for positioning and heading control of ships and offshore rigs,” IEEE Control Systems Magazine, vol. 29, no. 6, pp. 32-46, 2009.
- [5] D. K. Le, J. M. Kim, and T. K. Nam, “A study on the ship movement estimation by using Kalman filter,” The Conference of Navigation and Port Research, pp. 261-262, 2012 (in Korean).
- [6] S. Y. Hwang, J. Y. Heo, K. T. Hong, and J. H. Lee, “Time series data analysis and fault diagnosis of plant process equipment using statistical machine learning method,” Korean Journal of Computational Design and Engineering, vol. 23, no. 3, pp. 193-201, 2012 (in Korean).
- [7] D. A. Tobon-Mejia, K. Medjaher, N. Zerhouni, and G. Tripot, “A data-driven failure prognostics method based on mixture of Gaussians hidden Markov Models,” IEEE Transactions on Reliability, vol. 61, no. 2, pp. 491-503, 2012.
- [8] S. K. Kim and J. K. Hong, “Swell noise attenuation using a cascade of F-X filter and median filter,” Korean Society of Earth and Exploration Geophysicists, vol. 15, no. 4, pp. 199-208, 2012 (in Korean).
- [9] S. I. Ahn, S. J. Choi, and S. H. Kim, “Development of fault diagnostic algorithm based on spectrum analysis of acceleration signal for wind turbine system,” Journal of Korean Institute of Intelligent Systems, vol. 22, no. 6, pp. 675-680, 2012 (in Korean).
- [10] Y. J. Lee, D. Y. Kim, and M. S. Hwang, “A study on data pre-filtering methods for fault diagnosis,” Korean Journal of Computational Design and Engineering, vol. 2012, no. 2, pp. 301-308, 2012 (in Korean).
- [11] Y. W. Song, H. S. Lee, H. S. Park, Y. J. Kim, and

- J. Y. Jung, "A signal processing technique for predictive fault detection based on vibration data," The Journal of Society for e-Business Studies, vol. 23, no. 2, pp. 111-121, 2018 (in Korean).
- [12] K. T. Hong, J. Y. Heo, S. Y. Hwang, and J. H. Lee, "Analysis of sensor data for detecting the abnormal state of FD FAN," Korean Journal of Computational Design and Engineering, vol. 23, no. 2, pp. 137-143, 2018 (in Korean).
- [13] J. W. Lee, "A fuzzy neural-network algorithm for noisiness recognition of road images," Transactions of Korea Society of Automotive Engineers, vol. 10, no. 5, pp. 147-159, 2002 (in Korean).
- [14] O. S. Byun, "A study on the fuzzy recurrent neural networks for the image noise elimination filter," Journal of the Korea society of computer and information, vol. 16, no. 6, pp. 61-70, 2011 (in Korean).
- [15] K. Lee, S. I. Kang, J. W. Kwon, and S. M. Lee, "Speech enhancement using RNN phoneme based VAD," Journal of the Institute of Electronics and Information Engineers, vol. 54, no. 5, pp. 85-89, 2017 (in Korean).
- [16] H. J. Seo and J. H. Chang, "Integrated acoustic echo and background noise suppression based on deep neural network," The Journal of Korean Institute of Communications and Information Sciences, vol. 2017, no. 6, pp. 946-947, 2017 (in Korean).
- [17] G. K. Lee, Y. I. Kim, J. W. Lee, and W. R. Jo, "A design of adaptive neural network filter to remove the baseline wander of ECG," Journal of the Institute of Electronics Engineers of Korea, vol. 39, no. 1, pp. 76-84, 2002 (in Korean).
- [18] I. H. Cho, C. Y. Kwon, and J. S. Kim, "A study on noise cancellation of ECG signal using neural network," Journal of Advanced Information Technology and Convergence, vol. 6, no. 4, pp. 151-157, 2008 (in Korean).