

Global Report

텍스트 손상 묘사를 활용한 딥러닝 기반 지진 피해 건축물의 분류 연구

1. 서 론

1.1 2014년 나파 남부 캘리포니아 지진 개요

1.2 건축물 인벤토리 데이터 세트

1.3 회귀 신경망

1.4 LSTM 딥러닝 신경망

1.5 LSTM 기반 건축물 태그 표기 모델

2. 결과 및 토론

3. 결 론

1. 서 론

재난에 의해 영향을 받은 인프라시설물의 신속한 손상 평가는 비상 조치 및 복구 계획의 수립에 중요한 역할을 한다. 시설물 소유주, 사용자, 비상 관리자, 지자체 공무원은 신속 손상 평가를 통한 정보를 바탕으로 시기적절한 결정을 내린다. 지역 사회의 주요 기능을 갖는 시설물의 영향 때문에 건축물의 손상에 대한 시기적절한 정보는 재난 발생 사후 대응 계획 및 관리에 매우 중요하다. 건축물 손상의 범위와 관련하여 실시간적인 상황 인식의 부재로 재난에 대한 사회 반응이 효과적이지 않을 수 있다. 전통적인 지진에 의한 피해 조사는 개별 건축물에 대하여 엔지니어, 건축가 그리고 건설 사업 관련 다른 전문가들이 육안조사에 의존하였다. 육안 조사를 통해 손상된 건축물들은 ATC-20(Applied Technology Council)의 지침에 따라 분류하거나 태그를 붙인다. 건축물이 안전하고 즉시 재사용/거주가 가능한 경우 녹색 태그로 분류되고, 건축물이 불안전 하거나 붕괴된 경우에는 빨간색 태그로 분류한다. 건축물의 재사용 시 잠재적 위험이 있는 지 불분명할 경우 노란색 태그를 붙인다. 이러한 상세한 조사 과정은 지진에 의해 손상을 입은 건축물의 내진성능 관련 중요한 정보를 제공하기 때문에 반드시 필요하다. 그러나 현장조사과정은 수개월동안 지속되므로, 긴급 대응과 조기 복구 계획 수립이 불가능하다. 예를 들면 1994년 Northridge 지진 발생 후 건축물 손상관련 1차 조사는 2개월 이상 소요되었다.

기술의 발전은 지진 및 다른 극한 재난에 의한 인프라시설물에 대한 거의 즉각적인 정보를 생산할 수 있는 장치를 개발할 수 있게 되었다. USGS(United States Geological Survey)에 의해 개발된 ShakeMap과 ShakeCast는 무료 플랫폼으로, 몇 분 내에 지진 강도 및 그 지진에 대한 시설물 영향과 관련된 정보를 관계자들에게 전달한다. 해당 플랫폼은 지진 기록, 지진 특성과 손상된 시설물의 지진 취약성관련 정보를 바탕으로 평가한다. 스마트 장치와 소셜 미디어 플랫폼의 개발로 일반 대중에 의해 제공되는 정보를 이용한 재난 비상관리 관련 연구가 증가하고 있다. 최근 연구자는 샌디(Sandy) 허리케인에 의한 손상과 트위터(Twitter) 활동과의 밀접한 관련성을 보여주었다. 해당 연구는 건축물의 손상은 Hazus-Multi-Hazard의 모델을 통해 도출된 결과를 활용하여 손상을 정의하였으며, 신속한 손상 평가를 위한 소셜 미디어 플랫폼의 활용을 조사하였다. 또한 플로리다에서 발생한 매튜(Matthew) 허리케인이 발생한 기간 동안 포스팅된 손상 관련 트윗(tweet)을 도출하기 위하여 색인어 의미 해석법(Semantic Analysis)을 이용한 색인 사전(index dictionary)을 개발하였다. 여기서 의미 해석법이란 언어를 정보로 전환하는 방법을 의미한다. 해당 연구는 신속하게 주요 피해 지역을 발견하기 위하여 소셜 미디어 정보의 활용 가능성을 검증하기 위하여 소셜 미디어 기반 손상과 실제 손상 데이터와의 상관관계를 활용하였다.

본 연구는 텍스트 기반 손상 정보(texture damage description)를 활용하여 지진에 피해를 입은 건축물의 상태를 신속하게 평가하기 위한 딥러닝(Deep Learning) 기반의 프레임워크(Framework)의 적용에 대하여 조사하였다. 글로 표현된 손상 상태에 대한 묘사는 미디어의 리포트, 전문가의 조사 그리고 트위터(Twitter)와 페이스북(Facebook)과 같은 소셜 미디어 플랫폼을 통해 획득할 수 있다. 건축물의 손상 상태는 ATC-20 지침에 정의된 태그 관련 카테고리를 활용하여 분류하였다. 본 연구에서 개발된 프레임워크는 EERI(Earthquake Engineering Research Institute)로부터 획득한 2014년에 남부 나파지역의 캘리포니아 지진(2014 South Napa, California Earthquake) 데이터를 이용하여 검증하였다. 해당 글로 표현된 손상 정보는 지원자에 의해 작성된 3,423 건의 건축물 데이터 세트로 구성되어 있으며, 이 중 1,552건의 건축물은 녹색 태그로 분류되었고, 1,674건의 건축물은 노란색 태그 그리고 197건의 건축물은 빨간색으로 표기되었다. 해당 데이터 세트는 임의로 학습용과 검증용으로 분류하였으며, LSTM(long short-term memory) 기반 딥러닝 알고리즘을 구축하는데 활용되었다.

1.1 2014년 캘리포니아 남부 나파 지진 개요

2014년 8월 24일에 발생한 진도 6.0 규모의 캘리포니아 남부 나파 지진은 나파 도시의 건축물과 라이프라인(상,하수도 및 생활관로)에 손상을 발생시켰다. 최대 지진가속도(PGA, Peak Ground Acceleration)은 0.61g로 인명 손실은 비교적 크지 않았지만(300명의 부상자와 1명의 사망자), 2,000건

이상의 건축물이 보통에서 심각한 손상이 발생하였다. 경제적 피해액은 500만 불에서 1억 불로 추정된다.

주거형 건축물의 실질적 피해는 굴뚝과 크리플(Cripple) 벽체 손상에 집중되었다. 굴뚝 손상은 미세한 균열부터 벽돌의 탈락 그리고 붕괴까지 발생하였다. 몇몇 주거형 건축물에는 크리플 벽체의 기초 부분의 부분적인 혹은 전체 붕괴가 발생하였다. 조적조와 비연성 콘크리트 골조로 구성된 상업용 건축물은 구조적 손상이 발생하였다. 상업용 건축물의 비구조 요소의 손상으로 유리창이 깨지거나 스프링클러 파이프, 파사드(facade) 손상과 포도주 저장소(wineries)의 포도주 저장통(wine barrel) 보관대(stack)의 붕괴가 발생하였다. 라이프라인의 손상의 미미하였고, 손상된 시설물은 재난 발생 후 얼마 지나지 않아 안에 복구되었다.

1.2 건축물 인벤토리(Inventory) 데이터 세트

건축물의 인벤토리 데이터 세트는 손상에 대한 묘사, ATC-20 지침 기반 태그 분류 정보와 지진에 의해 영향을 받은 건축물의 위치와 특성으로 구성되어 있으며, EERI의 웹사이트에서 정보를 획득했다. 2014년 지진 발생 이후에 지진에 영향을 받은 건축물의 안전성을 평가하기 위하여 나파의 건축과는 건축가, 엔지니어 그리고 다른 건축물 관련 전문가들이 포함된 팀을 구성하였다. 해당 팀은 ATC-20의 태그 표기 절차를 바탕으로 건축물 입구에 플래카드를 설치하는 것으로, 건축물의 손상을 평가하였다. 그림 1은 데이터 세트에 포함된 3,423건의 건축물에 대한 ATC-20 태그가 표기된 공간 정보를 보여준다. 해당 태그 정보를 통해 도심지에 손상이 집중된 것을 볼 수 있지만, 전반적으로 손상이 도시 전체에 분산된 것을 볼 수 있다. 몇몇 건축물은 미가공 데이터 세트에 손상 묘사와 관련된 정보가 포함되어 있지 않으므로, 본 연구에서는 이와 관련된 건축물을 제외하였으며, 3,423건의 개별 건축물 중 45% 정도가 녹색, 49%가 노란색 그리고 6%가 빨간색 태그로 표기되어있다.

손상 상태에 대한 묘사는 한 개 혹은 두 개의 문장으로 구성되어 있으며, 손상 상태 평가는 소셜 미디어나 다른 소스를 통해 쉽게 얻을 수 있는 육안 조사 결과를 바탕으로 하였다. 다수의 기관에 의해 육안 조사가 이루어졌기 때문에 손상 묘사에 제시된 상세 수준은 상당히 다른 부분이 있으며, 사람과 건축물의 손상 범위에 따른 묘사에서 몇몇 단어들은 상당한 차이를 보인다. 같은 수준의 손상을 보인 경우에도 손상 묘사에 제시된 정보의 양이 다른 부분이 상당수 존재한다.

그림 2는 3,423건의 건축물에 대한 손상 상태의 묘사를 워드 클라우드(Word Cloud)로 나타낸 것이다. 해당 워드 클라우드는 손상 상태의 묘사에서 노출되는 주요 단어를 시각화 한 것으로 자주 노출된 단어는 크게 표기 하였다. 가장 빈번하게 노출된 단어는 굴뚝(Chimney)으로 그림에서 큰 폰트로 오렌지색으로 강조된 것이다. 특정 손상 유형과 지정된 태그와의 관련성을 평가하기 위한 예비 조사 단계로 단어 클라우드로 3개의 태그와 분류 작업을 수행하였다. 여기서 다수의 태그의 손상 묘사에서 중복으로 단어가 빈번히 발견된 경우 하나 이상의 태그에 해당 단어를 배치하였다. 그림 2의 워드 클라우드를 바탕으로 조사했을 때 다음과 같은 특징이 발견되었다. (1) 대부분의 건축물은 굴뚝 손상이 발생하였고, 노란색 태그로 자주 분류되었으며, (2) 많은 건축물이 미세한 균열이 발생하였고, 소수의 건축물은 붕괴가 발생하였거나 거주하기에 불안정한 것으로 조사되었다. (3) 구조체가 기울거나 기초부 또는 벽체에 손상이 발생한 경우 건축물은 빨간색으로 분류될 가능성이 높다. (4) 녹색 태그의 건축물들은 일반적으로 약간의 균열이 발생하였고, 구조적으로 양호한 상태를 보였다.

1.3 회귀 신경망(Recurrent Neural Network)

생물학적 신경계의 구성과 기능에 영감을 받은 신경망 모델은 신경 또는 절점 사이에 다수의 연결 관계로 구성되어 있다. 하나의 입력 신호는 다른 절점 또는 외부 자극으로부터 얻은 전체 정보의 집합체인 각 절점을 통해 받는다. 해당 신호는 출력 값 또는 다른 절점에 전달할 수 있는 출력 신호를 생산하는 활성화 혹은 전이 함수를 이용하여 처리된다. 이러한 신경망은 매우 효율적으로(신속) 많은 양의 정보를 처리할 수 있다.

신경망은 일반적으로 매우 복잡한 함수에 적용한다. x 라는 입력 세트가 있다면, 다층 구조의 신경망

은 임의로 영향인자를 샘플링하고, 각 절점의 전이함수를 통해 출력 값을 예측한다. 인공 신경망 모델을 학습시킬 때 예측된 출력 값(인공신경망을 통해 도출된 결과를 의미 함)은 실제 출력 값(실험 또는 학습에 사용된 출력 값을 의미함)과 비교를 하고, 그 오차는 역으로 주변의 가중치와 편향 값을 보정하는데 사용된다. 이러한 과정을 예측 오차의 값이 허용 오차 범위 내에 도달 할 때 까지 반복한다. 전통적인 인공신경망은 입력과 출력이 독립된 것으로 가정하여 글자 또는 문장과 같이 연속적인 출력 값을 예측하기에 부적합하다.

회귀 신경망(Recurrent Neural Network)은 연속적으로 현재 계산된 출력 값이 다음의 입력 값으로 사용되는 일련의 과정을 수행하기 위하여 개발되어 왔다. (즉, 기존과 달리 입력 값부터 출력 값까지 정보의 전달과정이 한 방향으로 흐르는 것이 아니라, 입력 정보를 통해 얻은 출력 값이 그 다음의 입력 값에 영향을 주는 방식이므로, 글자 또는 문장의 학습을 위해 활용할 수 있음) 그림 3은 전통적인 인공 신경망의 개념과 회귀 신경망의 방법을 비교한 것이다.

1.4 LSTM(Long Short-Term Memory) 딥러닝 신경망

LSTM은 회귀 신경망(Recurrent Neural Network)의 일종으로, 시제 연속(temporal sequence) 및 장기적인 의존성을 모델링하기 위한 능력은 전형적인 회귀 신경망에 비해 정확도 측면에서 우수하며, 여기에 추가적인 기계학습 기술을 도입하면 LSTM 기법을 글자 기반의 분류에 더 적합하게 만들 수 있다.

LSTM 신경망은 입력 층, 하나 혹은 여러 개의 은닉 층과 출력 층으로 구성되어 있다. LSTM 신경망은 각 은닉 층에 기억 셀(memory cell)이라고 부르는 특별한 유닛을 포함한다. 해당 기억 셀의 구성은 입력 게이트, 망각 게이트(forget gate)와 출력 게이트로 구성되며, 각 시간 별로 기억 셀의 상태를 세 개의 게이트로 결정한다. 입력 값은 입력 값의 특성과 이전 시간대의 기억 셀의 출력 값으로 구성된다. 망각 게이트는 셀의 상태를 파악하여 정보를 제외시킬지를 결정하며, 입력 게이트는 기억 셀에 추가될 정보를 명시하고, 출력 게이트는 정보를 출력 값으로 내보낼지를 판단한다.

1.5 LSTM 기반 건축물 태그 표기 모델

3,423건의 건축물 데이터 세트는 임의로 학습과 검증 세트로 구분하였다. 학습 세트(총 데이터의 75%)는 LSTM 예측 모델을 개발하는데 활용하였으며, 검증 세트(총 데이터의 25%)는 해당 모델의 성능을 평가하기 위하여 사용하였다. 이러한 임의로 구분된 데이터는 오버핏(overfit)에 대한 잠재성을 감지하고 알 수 없는 데이터 세트가 제공되었을 때 예측 능력을 평가할 수 있다. 여기서 오버핏은 주어진 데이터 세트에 결과가 과도하게 순응하는 것을 의미한다. 해당 문제는 인공신경망 학습에서 자주 생기는 오류로 주어진 데이터 세트에 대해서만 합리적인 결과를 도출하게 된다.

주어진 손상 태그와 건축물 손상 묘사는 LSTM의 입력 값으로 활용하기 위하여 워드 벡터(word vector)로 전환하였다. 각 워드 벡터의 길이(즉, 태그 관련 설명에 대한 단어 수)는 30개로 제한(데이터 세트의 최대 단어 수가 30개로 구성됨)되었으며, 손상 묘사가 30개 단어 보다 적은 경우 30개 단어의 길이와 동일하도록 여백을 만들었다. 이는 LSTM 모델이 동일한 단어 수로 구성된 입력 데이터로 구성되었을 때 보다 우수한 성능을 보이기 때문이다.

민감도 분석(Sensitivity Study) 결과를 토대로 150개로 구성된 은닉 층과 빨간색, 노란색, 녹색 태그의 출력 결과를 예측할 수 있는 3개의 완전 연결 층을 활용하여 LSTM 모델을 개발하였다. 0.01의 학습율을 갖는 예측 모델을 개발하기 위하여 경사 경계 알고리즘(gradient threshold algorithm)은 Adam 최적화 모델을 사용하였다. 검증 데이터의 예측 정확도와 150개 은닉 층의 반복계산의 관계를 살펴봤을 때 150번의 반복계산 후 정확도(전체 태그 중 정확히 예측한 비율)가 일정한 값을 보였다.

그림 9는 개략적인 LSTM 기반의 손상 분류 예측 모델을 나타낸 것으로, ‘Minor cracks in exterior wall’과 같이 5개의 단어로 구성된 손상 묘사의 단어 수는 ‘0’을 추가하는 방식으로 30개의 단어로 전환하였다. 해당 단어 벡터는 각 단어의 중요성을 파악하기 위한 단어 삽입 후 LSTM 셀로 전달된다. 반복계산 완료 후 최신 LSTM의 출력 값은 완전 연결 층으로 전달되고, 해당 연결 층의 가중치와 편향과 연계하여 손상상태를 정의한다. 여기서 주의할 점은 LSTM 모델의 학습 중에 LSTM의 셀과 출력 층의 가

중치와 편향은 실제 예측과 관측된 예측 결과의 차이를 바탕으로 산정된 것이다.

2. 결과 및 토론

개발된 LSTM 모델의 예측 성능은 오차 행렬을 이용하여 평가한다. 오차 행렬은 관측된 태그 표기와 예측된 태그 표기와의 관계를 나타낸 것이다. 그림 5의 (a)와 (b)는 학습과 검증 데이터 세트를 기반으로 표기된 태그의 오차 행렬을 보여준다. 태그에 지정된 LSTM 모델의 성능은 정밀도와 재현율(Recall)을 통해 평가할 수 있다. LSTM 알고리즘을 이용하여 정확하게 표기된 예측물은 오차 행렬의 4번째 행에 제시되었으며, 기계 학습 알고리즘을 통해 정확하게 지정된 실제 태그 예측물은 재현율로서 4번째 열에 나타났다. 높은 재현율과 정밀도는 LSTM이 얼마나 정확히 태그를 예측하는지 보여준다. 전체 태그 수 대비 정확히 예측된 태그의 수에 대한 비율은 예측 모델의 정확도를 나타낸 것으로 정의하였다.

본 연구에서 개발된 LSTM 모델은 학습 데이터 세트에서 녹색 태그에 대하여 98%의 정밀도를 보였다. 이는 1,167개의 녹색 태그 예측이 실제 녹색 태그의 1,148개와 일치하는 것을 보여준다. 또한 LSTM은 1,164개의 실제 녹색 태그 중 1,148개를 정확히 표기했으며, 이는 99%의 재현율을 보여준다. LSTM 모델은 학습용 데이터 세트에 대하여 전반적으로 98%의 높은 정확도를 보였다. 검증용 데이터 세트(학습에 사용되지 않은 데이터 세트)는 86%의 정확도를 보였다. 따라서 개발된 LSTM 모델은 86%의 정확도로 손상에 대한 텍스트 묘사를 기반으로 ATC-20의 태그 표기를 하는 것으로 나타났다. 해당 모델은 녹색과 노란색 태그의 표기는 우수한 것으로 나타났지만, 빨간색 태그의 표기는 낮은 정확도와 재현율을 보였다. 이는 데이터 세트에서 상대적으로 빨간색으로 표기된 건축물이 부족하다. 또한 데이터의 비대칭적인 분포는 손상 상태의 묘사에 대한 세부 수준에 따라 데이터에 대한 예측 능력이 과소평가 되는 경향이 있다. ‘*현관의 기초에서 이동이 발생한 건축물 (Building shifted on foundation of front porch)*’의 손상이 발생한 경우 건축물이 불안정하므로 빨간색 태그로 설정해야 하지만, 현장 조사에서 조사자는 노란색 태그로 표기하였다.[학습 데이터 세트에서 LSTM 모델에 의해 표기된 유사 손상을 기반으로 LSTM은 해당 손상 상태를 빨간색 태그로 표기하였다. 또한 균열 관련 애매한 설명은 LSTM 예측 모델에 혼란을 발생시킨다. 예를 들면, ‘*문의 모서리와 벽체에 균열 발생 (Cracks at door corners and walls)*’과 같은 손상 묘사는 LSTM 모델에서 빨간색으로 태그가 되지만 실제 태그는 녹색으로 표기되었다. 데이터 세트를 바탕으로 작성된 태그 별 단어 클라우드를 살펴보면 ‘균열’이라는 단어는 주로 녹색 태그로 분류되고 있고, ‘벽체 (wall)’라는 단어는 빨간색 태그로 분류가 된다는 점을 주목할 필요가 있다. 이러한 몇 가지 결함에도 불구하고, LSTM 모델은 합리적으로 ATC-20 지침의 태그를 잘 묘사하는 것으로 나타났다.

데이터 세트의 사용 유형과 관련해서 앞에서 언급한 것과 같이 임의로 학습을 위해 75%의 데이터 세트가 선택 되었고, 25%의 데이터 세트는 검증을 위하여 사용되었다. 그림 5의 오차 행렬의 결과는 75%의 학습 데이터와 25%의 검증용 데이터를 바탕으로 나타난 것으로, 학습과 검증 데이터의 크기(양)를 달리하여 모델 정밀도 관련 민감도 해석(Sensitivity Study)을 수행했다. 검증 데이터 세트의 비율을 10%에서 50%까지 증가시키면서, 모델의 민감도 해석을 수행했을 때 데이터 세트의 비율에 따른 모델의 정확도의 차이는 5% 이내로 나타났다.

LSTM 모델로 예측된 결과와 실제 결과를 각 구성 요소(녹색, 노란색, 빨간색 태그)별로 나누어 살펴보면, 녹색 태그로 지정된 건축물은 대부분 미세한 균열과 관련되어있고, LSTM의 예측 정밀도는 78%로 나타났다. 해당 LSTM 모델은 노란색 태그로 지정된 건축물 중 247개의 굴뚝 손상 중 244개를 정확하게 예측하였으며, 이는 오차율이 1.2% 미만임을 의미한다. 대부분의 빨간색 태그로 지정된 건축물은 기초부와 건축물이 기울어지는 손상과 관련되었으며 이러한 손상의 LSTM 모델의 예측 오차는 15% 미만으로 조사되었다.

3. 결론

지진 발생 이후의 건축물 손상 평가는 긴급 대응 및 복구 계획을 위한 결정에 정보를 제공한다. 건축물의 손상 상태는 미세한 균열부터 붕괴까지 매우 다양할 수 있으며, 충격이 가해진 지역 분포는 개별 건물마다 사람에 의해 평가하는 복잡하고 시간이 많이 소모되는 과정으로, 손상 크기와 조사관에 따라 수개월 동안 작업이 이루어질 수 있다. 본 연구는 텍스트 기반 손상 묘사를 바탕으로 지진에 의해 손상된 건축물의 신속한 평가를 수행하기 위하여 딥러닝 기반의 LSTM 신경망을 적용하였다. 해당 방법론은 2014년 캘리포니아 남부 나파 지진에 의해 손상된 건축물의 손상 상태 묘사를 통해 검증하였다. 본 연구를 통해 개발된 LSTM 모델은 전반적으로 86%의 정확도를 보였으며, 특히 검증용 데이터 세트에 대하여 녹색 태그는 94% 그리고 노란색 태그는 84%의 예측 정밀도를 보였다. 그러나 빨간색 태그에 대한 LSTM 모델의 정밀도는 59%로 조사되었다. 이는 빨간색으로 분류된 건축물의 수가 적으며(학습에 활용된 데이터 수가 적음을 의미), 손상 묘사에 제시된 세부 사항의 차이가 크기 때문이다.

여기서 주의할 점은 본 연구에서 제안된 LSTM 기반의 건축물 손상 분류 관련 프레임워크가 사람이 직접 수행하는 ATC-20 지침의 손상 평가를 대신할 수 있다는 것이 아니라 장기적인 관점에서 해당 방법론이 실시간 정보를 생산하는 것을 통해 사람이 직접 수행하는 태그를 표기하는 과정을 보완할 수 있다는 점이다. 예를 들면 모바일 앱을 통해 건물주와 사용자가 지진이 발생한 직후에 해당 건축물의 손상에 대한 설명을 웹 기반의 데이터베이스에 업로드하고, 이를 LSTM 알고리즘의 입력 값으로 활용할 수 있다. 건축물의 특정 손상으로 분류(출력 값으로 녹색, 노란색, 빨간색 태그를 의미)된 정보를 긴급 대응 및 복구 계획을 수립하는 사람에게 지진발생 직후 제공할 수 있다.

LSTM 기반의 방법론의 전반적 가능성은 검증되었지만, 해당 방법론의 적용 범위를 넓히기 전에 현재 연구는 몇 가지 제한 사항을 포함하고 있다. 예측 모델은 하나의 지진에 대한 건축물 피해 관련 데이터 세트를 바탕으로 개발 및 평가되었으며, 해당 지진은 다른 지진에 비해 중규모 수준이다. 또한 2014년에 발생한 캘리포니아 남부 나파 지진의 피해는 대부분 주거형 건축물에 발생하였고 그 구성 요소의 유형도 제한적이다. 추후 연구를 통해 지진 규모에 따라 다양한 지진과 건축물 손상 유형을 고려하여 모델을 학습하고, 평가해야 한다.

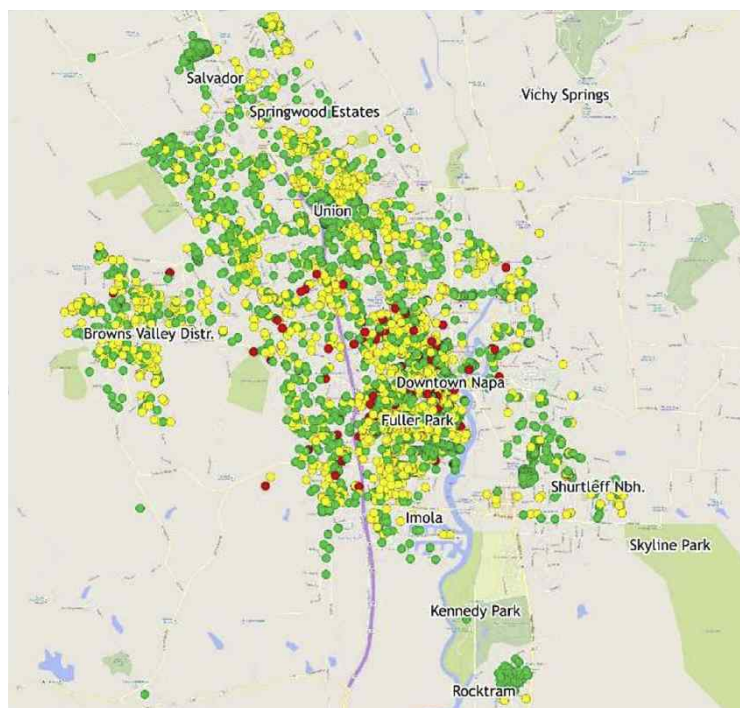
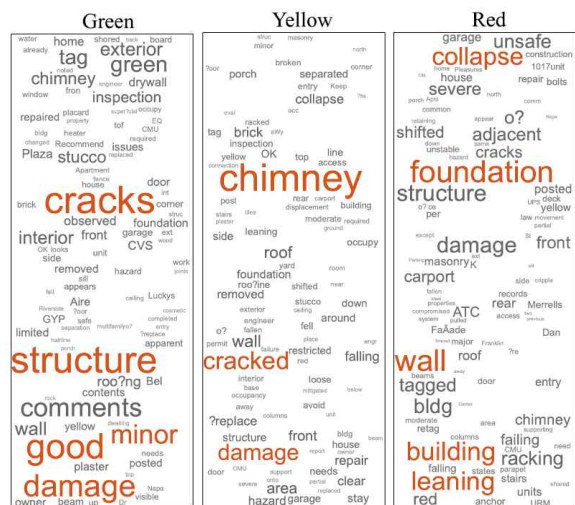


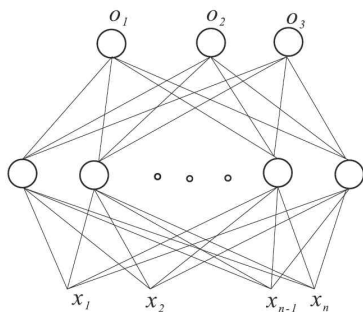
그림 1. ATC-20 손상 태그가 표기된 공간 정보



(a) 전체 워드클라우드

(b) 태그 별 워드 클라우드

그림 2. 손상 상태 묘사 관련 데이터 세트의 워드 클라우드

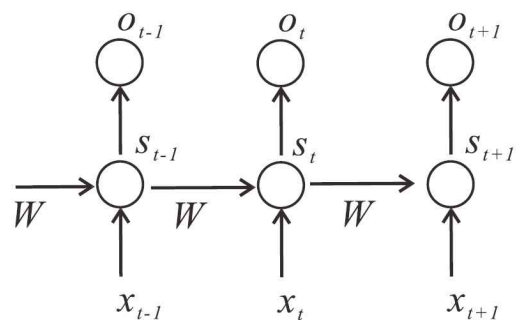


Neural network



Idealized representation
of neural network

(a) 일반 인공신경망 구조



(b) 회귀 신경망 구조

그림 3. 인공지능경망 구조 비교

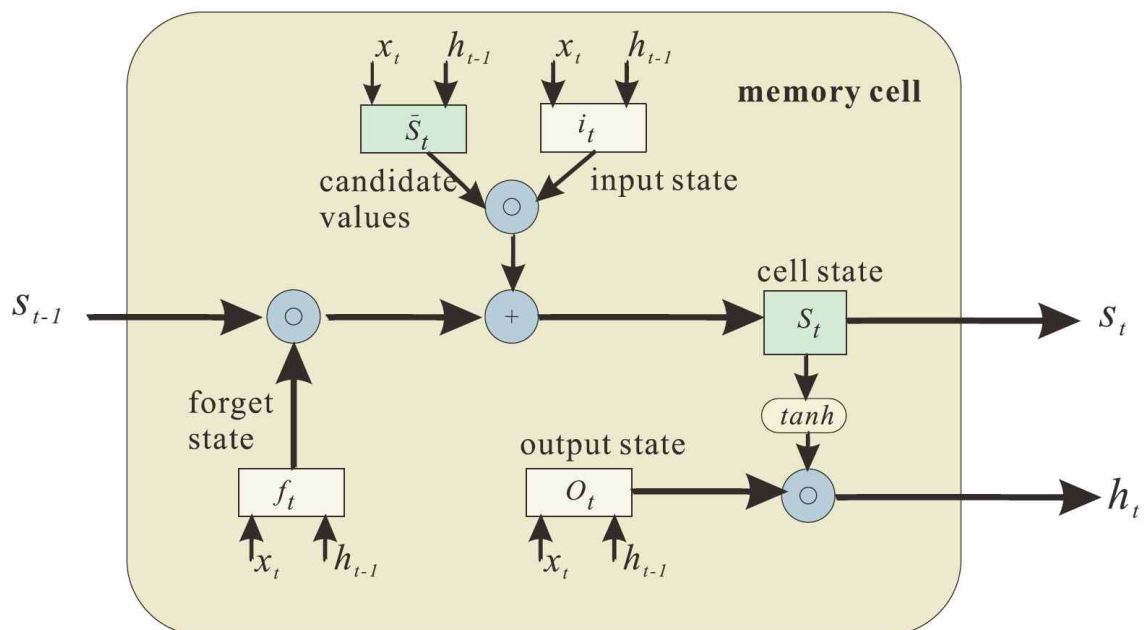


그림 4. LSTM 신경망 모델의 기억셀 구조

		Predicted class			
		G	Y	R	
Actual class	G	1148 (44.7%)	13 (0.5%)	3 (0.1%)	99%
	Y	16 (0.6%)	1239 (48.2%)	1 (0.0%)	99%
	R	3 (0.1%)	3 (0.1%)	142 (5.5%)	96%
		98%	99%	97%	98%

(a) 학습 데이터 세트

		Predicted class			
		G	Y	R	
Actual class	G	325 (38.0%)	59 (6.9%)	4 (0.4%)	84%
	Y	18 (2.1%)	382 (44.7%)	18 (2.1%)	91%
	R	4 (0.4%)	14 (1.6%)	31 (3.6%)	63%
		94%	84%	59%	86%

(b) 검증 데이터 세트

그림 5. LSTM 모델 기반 오차 행렬

<원문제목> Deep learning-based classification of earthquake-impacted buildings using textual damage descriptions

<원문출처> International Journal of Disaster Risk Reduction, ELSEVIER