

랜덤 포레스트 기반 위험도 지수 평가항목 타당성 확인에 따른 터널 안전관리 정책지원 방안

정소영

동양대학교 스마트안전시스템학부

Supporting Road Tunnel Safety Management Policies by Random Forest Based Risk Factor Validation

Soyoung Jung

School of Smart Safety System, Dongyang University

요약 본 연구는 교통안전 측면에서 정부의 터널 안전관리 강화정책을 계량적으로 뒷받침하기 위한 첫 단계로서, 터널의 위험도 등급산정을 위한 기존 위험도 지수 세부평가항목들의 타당성 확인에 따라 방재 시설 보강이 필요한 터널 및 해당 방재 시설 목록선정을 목적으로 하였다. 이러한 연구목적을 위하여 국내 고속도로 전체 터널에서 발생한 교통사고 데이터를 기반으로 많은 데이터 필드 중 부상 사고에 영향을 주는 주요 위험요인들을 랜덤 포레스트 알고리즘을 이용, 우선으로 추출하고 기존 터널 위험도 지수 세부평가항목들과 비교하였다. 이어 랜덤 포레스트와 분산팽창요인 기반 이분형 로지스틱 리그레션 모델을 비교하여 최적 모델을 선택하고, 최적 모델을 이용한 부상 사고 발생확률 예측에 따라 보강할 방재 시설 종류와 해당 터널의 순위화를 시행하였다. 그 결과, 기존의 터널 위험도 지수 세부평가항목으로서 타당성이 재확인된 위험요인들은 주행거리계, 터널 높이, 커브, 대형차 관련 항목으로 나타났다. 또한, 27개의 고속도로 터널에서 방재 시설 보강이 필요하며 특히 설치보강이 필요한 방재 시설로는 경보 혹은 모니터링 관련 설비인 것으로 확인되었다.

Abstract This study aimed to quantitatively validate the strategy specifications of the recent government policy in its initial stage. In particular, the current study used the random forest algorithm to identify major risk factors, among many data fields, that significantly affected freeway tunnel crash severity. Further, this study compared the random forest-based risk factor identification with the existing list of risk factors for disaster prevention facility installation in tunnels. This study also identified the types of disaster prevention facilities needed for tunnel safety improvement and selected freeway tunnels with the relevant disaster prevention facilities. This is achieved by comparing the random forest-based binary logistic regression with the variance inflation factor-based regression. Based on the analysis, vehicle kilometer traveled, tunnel height, horizontal curve section, heavy vehicle were confirmed as risk factors affecting injury involved crashes in a freeway tunnel. Especially, communication or monitoring facilities were significantly identified to decrease tunnel crash severity for disaster prevention, and 27 freeway tunnels needed warning facilities.

Keywords : Tunnel, Safety Improvement, Risk Factor, Disaster Prevention Facility, Random Forest

*Corresponding Author : Soyoung Jung(Dongyang Univ.)

email: jung2@dyu.ac.kr

Received October 7, 2021

Accepted February 4, 2022

Revised November 15, 2021

Published February 28, 2022

1. 서론

1.1 연구의 배경 및 필요성

국토 면적의 약 70%가 산지인 국내의 지형적 상황에서 도로 터널(이하 터널)은 이동의 편리성과 효율성을 향상해 왔다. 그러나 터널은 폐쇄적인 구조로 인해 시야 확보가 어렵고 위험회피 공간이 제한되어 있으므로 터널 건설이 늘어날수록 관련 교통사고도 증가하고 있다. 이에 2020년 5월, 국토교통부는 도로 인프라 개선을 중심으로 도로교통 안전강화 특별대책을 발표하였는데 관련 중점과제 중 하나로 터널 안전관리 강화를 선정하고 관련 세부내용으로 1) 전 터널 관리등급 재산정 및 2) 터널에 대한 방재 시설의 설치와 보강을 발표하였다[1]. 현재 국내의 터널 관리등급은 도로 터널 방재 시설 설치 및 관리지침에 의하여 터널 연장(연장등급) 및 위험도 지수(위험도 등급)에 따라 나누어지는데 특히 위험도 지수 산정 시 고려해야 할 세부평가항목들은 사고확률, 터널특성, 대형차량, 정체 정도, 통행방식과 관련된 위험요인 목록들로 구성되어 있다[2]. 또한, 이러한 터널의 연장 및 위험도 등급에 따라 소화 및 소화 활동, 경보, 대피, 비상 전원설비와 같은 방재 시설들도 상이하게 설치하도록 규정되어 있다[2].

그러나, 현재의 터널 관리등급 기준은 2009년 개정된 이후 터널 주변 교통환경 변화 때문에 재조사의 필요성이 대두되고 있다. 특히, 터널 위험도 등급산정에 활용되는 위험도 지수 세부평가항목 중 하나인 사고확률은 터널 구간의 연평균 일 교통량과 터널 연장의 곱으로만 계산하도록 규정되어 다양한 위험요인들의 영향을 반영한 사고확률 예측치를 고려하지 않고 있다. 이에 교통환경 변화를 반영하여 터널 위험도 지수 산정의 세부평가항목들을 계량적으로 확인하고 범용성 있는 위험요인 목록으로 수정할 필요가 있다. 또한, 터널 내부에서 일어난 교통사고의 경우 사고 직후에는 부상자의 부상 상태가 가변했다 할지라도 터널의 폐쇄적인 공간구조로 인해 구조가 늦어져 사고 이후 부상자가 중상이나 사망으로 이어질 수 있으므로 터널 방재 시설의 일차적인 목적은 사고 후 피해, 즉 터널 사고 심각도의 완화라고 할 수 있다. 따라서, 터널 방재 시설의 설치 및 보강을 위해서는 터널 사고의 심각도와 관련 지어 고려하지 않을 수 없다.

국토교통부 도로 교량 및 터널 현황조사에 따르면 2019년 전국 터널은 총 2,682개소, 총연장 2,076.5km로 2010년 대비 각각 2배 이상 증가하였는데, 특히, 고속도로 터널은 2019년 국내 전체 도로 터널 개소 및 총

연장 중 각각 약 45% (1,204개소) 및 52% (1,079km)를 차지하고 있다[3]. 더불어 도로교통공단 교통사고분석시스템 통계에 따르면 터널 건설과 연장이 늘어남과 따라 터널 사고 및 사망 포함 부상자 수도 2016년 이후 매년 증가하고 있는데, 고속도로 터널 내 부상 사고는 2013년 100건에서 매년 증가하여 2017년 161건으로 2013년 대비 61%가 증가한 실정이다. 그런데도 도로 터널의 대표 격인 고속도로 터널 전체에서 발생한 교통사고 심각도를 고려하여 정부의 터널 안전관리 강화정책을 계량적으로 뒷받침하는 국내연구가 아직 없었다.

1.2 연구의 목적

본 연구는 앞에서 언급한 대로 정부가 발표한 터널 안전관리 강화를 위한 두 가지 세부정책 추진을 뒷받침하기 위한 첫 단계로서, 1) 터널 위험도 등급 재산정을 위한 위험도 지수 세부평가항목들의 타당성 확인과 2) 주요 방재 시설 및 해당 방재 시설을 설치해야 할 터널목록을 선정하고자 한다. 이와 같은 연구목적을 달성하기 위하여 본 연구는 2012년부터 2016년까지 국내 고속도로 전체 터널에서 발생한 교통사고 데이터와 터널구조 및 실시간 교통류 정보를 결합하였다. 이어 랜덤 포레스트(Random Forest, 이하 RF) 알고리즘 및 분산팽창요인(Variance Inflation Factor, 이하 VIF) 기반 이분형 로지스틱 리그레션(Binary Logistic Regression, 이하 BLR) 모형 비교를 통하여 고속도로 터널 교통사고 심각도에 영향을 끼치는 주요 위험요인 목록을 추출하였다. 또한, 터널 교통사고 심각도 확률예측 모형 결과에 따라 가장 먼저 보강해야 할 방재 시설 종류와 이러한 방재 시설의 설치가 필요한 터널들도 확인하였다.

2. 선행연구

2.1 터널 위험요인 선정 및 방재 시설

서론에서 언급한 연구목적에 따라 교통사고 심각도에 영향을 미치는 변수선정 방법론과 터널 위험요인 및 방재 시설에 초점을 맞추어 다음과 같이 선행연구들을 살펴해보았다. 우선, 교통사고 심각도에 영향을 미치는 위험요인으로 독립변수를 추출한 선행연구에서는 전통적 방식인 회귀모형을 비롯하여 결정트리(Decision Tree, 이하 DT), 서포트 벡터 머신(Support Vector Machine, 이하 SVM), RF와 같은 데이터의 분류기법들이 쓰였다.

최근 다항로지트 및 RF 모델을 비교하여 서로 다른 데이터 형태 (연속형 혹은 명목형)로 이루어진 변수 중에서 교통사고 심각도에 유의미한 영향을 주는 주요 위험요인을 추출한 연구에서 RF 모형화가 다중공선성 (multicollinearity)을 줄이고 변수들의 비선형적인 영향을 잘 추출하여 교통사고 심각도를 보다 정확하게 예측할 뿐 만 아니라 다항로지트와 같은 회귀모형보다 데이터에 덜 민감한 것으로 확인되었다[4]. BLR, DT, RF 모델을 비교하여 심각한 부상 사고에 영향을 미치는 요인들을 예측한 연구 및 가용 데이터가 대부분 범주형 형태일 경우 교통사고 차량 운전자의 부상 심각도 영향요인 추출을 위하여 BLR, SVM, RF와 결합한 DT의 세 모형을 비교, 분석한 연구에서도 RF 혹은 RF 결합 DT 모델이 예측의 정확성 측면에서 타 모델보다 더 우수한 것으로 확인되었다[5,6].

다음으로, 터널 교통사고 위험요인에 관한 선행연구에서는 제한속도, 터널구조, 노면 상태, 운전자 요인 및 이러한 요인들의 인터랙션 등 다양한 형태의 위험요인들이 확인되었다. 그 예로 터널과 같은 지하시스템 변수 간의 상호의존성을 고려하여 정량화할 수 있는 베이지안 네트워크 기반 정량적 위험평가 방법론의 구조를 소개한 연구에서 도로 터널의 베이지안 네트워크 기반 사고위험 모형화에 고려한 위험요인 지표는 교통량, 출입구, 위험물 운반 차량 비율, 단/양방향 교통, 터널 조명, 터널 곡선반경, 터널 경사, 차선 수, 차선폭, 서비스 수준, 차선 이동, 터널 내 속도제한으로 매우 다양한 요인들이 포함되었다[7]. 터널의 존재와 지형의 변화로 위험 구간이 연속되는 미시령 관통 도로에서 구간속도위반 단속 장비 설치 효과를 분석한 연구에서는 터널 입구부터 요금소 전방의 전 구간에 구간속도위반 단속 장비를 설치하였을 때 설치 전 대비 월간 교통사고 건수가 약 46% 감소하였으나 철거 후에는 설치 전보다 월간 교통사고 건수가 약 65% 증가한 것으로 나타나 터널 구간에서의 통행속도와 교통사고의 연관성을 보여주었다[8]. 비슷한 연구로 시물레이션 기반 터널 부 가변형 제한속도 (Variable Speed Limit, 이하 VSL) 표출제어 알고리즘 개발연구에서도 터널 부 VSL 표출제어의 경우, 구간별로 통행속도가 감소할 뿐만 아니라 정체 시 교통류와 정상류 간의 속도 차이가 감소하는 것을 확인하여 터널 부 교통사고에 교통류 속도제어 시스템의 효과가 큼이 확인되었다[9].

인적 요소와 관련해서는 노면의 상태와 운전자의 안전거리 확보 의지가 겨울철 대형 터널 사고에서 주요 요인이 확인되었을 뿐만 아니라, 법규준수 의지 요인과 터

널 내 정보인지 요소들이 대피 행동의 첫 단계인 위험경고 신호 인식과 유의미한 관계가 있는 요인임으로 나타났다[10,11]. 특히, 운전자의 피로나 과속과 같은 운전자 요인이 커브나 내리막과 같은 지형변화나 오전 혹은 저녁 시간과 같은 시간 요인과 결합하였을 때 중상이나 산악지역의 고속도로 터널에서의 사망사고 확률을 높이는 것으로 나타났다[12].

끝으로, 터널 방재설비와 관련한 연구에서는 터널의 폐쇄적인 특징으로 화재와 같은 2차 피해 저감 대책에 관한 연구들이 주요 흐름인 것으로 나타났다. 터널 방재 시설 위험도 감소 효과를 분석하여 터널 화재 시 위험도를 낮출 수 있는 인자들에 대한 순위를 제시한 연구에서 방재 시설별 비용대비 사람의 부상이나 사망 발생에 대한 위험도 감소 효과 면에서 우선순위는 송수관설비, 비상조명, 대피 통로, 제연설비의 순으로 확인되었다[13]. 또한, 도로 터널 내부에서 교통사고 및 정전 혹은 차량화재 발생 시 방재설비들의 안전관리 방안을 분석한 연구들에서 사고 이후 터널 내 2차 재난 상황을 막기 위하여 CCTV 설치 및 관련 전원 이중화, 적정 지름과 유량을 갖춘 제트 팬이 터널 내 사고 발생 시 사고원인의 확인 및 신속한 제연을 위한 중요 방재설비임도 확인되었다 [14,15].

2.2 기존연구와의 차별성

본 연구는 고속도로 터널 전체에서 발생한 교통사고의 데이터 세트에 기록되지 않은 터널구조 특성 및 사고 발생 터널의 방재 시설 및 실시간 교통류 데이터를 교통사고와 매칭하여 통합된 터널 교통사고 데이터를 구축하고, 이러한 통합 데이터를 기반으로 정부의 터널 안전관리 강화정책을 교통안전 측면에서 계량적으로 지원하는 국내 최초의 연구이다. 현재 국내의 기존의 터널 등급 평가 기준을 정립한 연구들에서 제시된 터널 위험도 지수 세부평가항목은 국외의 세부평가항목을 그대로 사용한 것으로 국내 터널 사고데이터를 기반으로 계량적으로 추출된 세부 평가항목이 아니다[16,17]. 특히, 터널 방재 시설 설치의 일차적 목표는 사고 후 심각도 완화인 데 반해 교통사고 심각도 확률예측 기법을 정부의 터널 관리등급 재산정 및 방재 시설 보강과 연결하여 분석한 국내 연구는 아직 없는데 본 연구는 고속도로 터널에서 발생하는 교통사고 심각도 확률예측에 기반, 터널 관리등급 재산정에 따른 위험요인 추출 및 방재 시설의 보강을 위한 계량적 방법론을 제시한 차별성 있다.

3. 데이터

3.1 터널 사고의 정의

터널 주변은 아래 그림에서 보는 바와 같이 구분할 수 있는데 그중 터널 입구와 출구 주변의 개방된 구역을 각각 터널 진입 구역 (Access Zone, 이하 AZ) 및 진출 구역 (Exit Impact Zone, 이하 EIZ)으로 부른다[18]. 본 연구에서 터널 사고는 터널 진입 구역으로부터 진출 구역까지 차량에 의하여 발생한 교통사고만을 고려하였으며, 이러한 교통사고를 본 연구에서 터널 사고로 정의하였다.

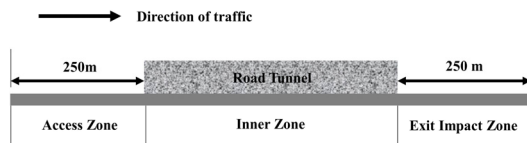


Fig. 1. Tunnel zone division

또한, 터널 사고 관련 선행연구들에서 진입 구역 및 진출 구역의 길이는 100m에서 250m 정도까지 고려하였다[19,20]. 이러한 선행연구 결과를 고려하고, 터널 입구 혹은 출구에서 더 넓은 범위의 주변 상황을 반영하기 위하여 본 연구에서는 진입 구역 및 진출 구역의 길이를 250m로 지정하였다.

3.2 기존 터널 등급 구분

방재 시설 설치를 위한 국내의 터널 위험도 등급은 터널의 위험도 지수 세부평가항목별 위험도 지수 값의 합 (이하 X)을 기준으로 구분하며, 터널 방재 시설은 터널의 연장 혹은 위험도 등급에 따라 달리 설치하도록 규정되어 있는데 터널 방재 시설로는 소화, 경보, 피난 대피, 소화 활동, 비상 전원설비가 포함된다[2]. 참고로 터널의 위험도 등급산정을 위한 X값의 범위에 대한 기준은 Table 1과 같다.

Table 1. Current standard of tunnel class

Class	Tunnel length (L)	Sum of risk indices (X)
1	$L \geq 3000\text{m}$	$X > 29$
2	$1000\text{m} \leq L < 3000\text{m}$	$19 < X \leq 29$
3	$500\text{m} \leq L < 1000\text{m}$	$14 < X \leq 19$
4	$L < 500\text{m}$	$X \leq 14$

국토교통부 도로 터널 방재 시설 설치 및 관리지침 <표. 2.2>에 따르면 Table 1에서 X값은 터널 사고에 영향을 미치는 주요 위험요인인 세부평가항목을 선정, 각 세부평가항목의 값을 범주화하고 범주별로 상이하게 부여된 위험도 지수 값을 더하여 계산한다. 터널 위험도 등급 구분을 위한 위험도 지수 세부평가항목은 다음과 같다[2]: 1) 사고확률 관련 세부평가항목: 범주화된 연평균 일 교통량과 터널 연장의 곱 (이하 주행거리계); 2) 터널 특성 관련 세부평가항목: 범주화된 진입부 경사도, 입/출구 표고 차, 터널 높이, 곡선반경; 3) 대형차량 관련 세부평가항목: 범주화된 대형차 혼입률, 대형차 주행거리계, 위험물 수송차에 대한 감시시스템 유무, 위험물 수송에 대한 유도시스템 유무; 4) 정체 정도 관련 세부평가항목: 범주화된 서비스 수준 (Level of Service, 이하 LOS), 터널 내 합류/분류 여부, 터널 전방 교차로/신호등/요금소 유무; 5) 통행방식 관련 세부평가항목: 대면통행 여부, 일방통행 여부.

본 연구목적에 따라 위에서 나열한 터널 위험도 지수 세부평가항목들의 범주별로 지정된 위험도 지수 수치는 따로 제시하지 않았다. 또한, 본 연구에서는 고속도로 터널이 대상이므로 일방통행만 고려되었고 LOS를 제외한 정체 정도 관련 세부평가항목은 고속도로 터널 상황에서는 해당 사항이 없어 고려하지 않았다. 더불어, 터널 내부의 위험물 수송 차량에 대한 감시시스템 및 유도시스템 유무와 터널 입/출구 표고 차에 대한 데이터는 수집의 한계가 있어 본 연구에서는 기존의 터널 위험도 지수 세부평가항목 리스트 중 최종적으로 터널 주행거리계, 터널 진입부 경사도, 터널 높이, 곡선반경, 대형차 혼입률, 대형차 주행거리, LOS의 타당성을 확인하였다.

3.3 분석 데이터

본 연구에서는 한국도로공사가 제공한 2012년부터 2016년까지 국내 고속도로에서 발생한 전체 교통사고 중 앞서 정의한 터널 교통사고만을 고려하였다. 그 결과, 439개의 터널에서 발생한 교통사고 케이스가 총 1210건이 수집되었는데 그 중 사망사고가 28건, 중경상을 포함한 부상 사고가 196건, 대물 피해사고가 1014건으로 나타났다. 터널 교통사고 순간에는 사망이나 중상을 포함하지 않은 가벼운 부상 사고라 할지라도 터널의 폐쇄적인 공간구조로 인해 대응이 늦어져 사고 이후 결과가 심각해질 우려가 있으므로 본 연구에서는 터널 사고의 심각도를 사망, 중경상을 포함한 부상 사고와 차량 탑승자의 부상이 없는 대물 피해사고, 즉 두 개의 레벨로 나누

어 모형을 실시하였다.

한국도로공사가 제공한 고속도로 교통사고 데이터 세트에는 종속변수로 사용될 사고 심각도 이외에도 각 사고 케이스 별 발생 위치 및 시간, 도로환경, 사고형태, 교통흐름, 차량 및 운전자 등과 관련된 많은 정보가 기록되어 있으나, 본 연구에서 사용한 터널 사고 케이스의 경우에는 해당 터널의 기하구조와 방재 시설 현황정보가 기록되어 있지 않았다. 특히, Fig. 2에서 보는 바와 같이 터널 방재 시설 종류에는 CCTV, 제연, 소방, 비상발전설비, 대피 통로, 정보표시판, 비상방송 설비, 사고 자동감지 설비가 포함되어 있다[21].



Fig. 2. Disaster prevention facilities in tunnel[21]

이에, 각 터널 사고 일시와 이정을 기반으로 해당 터널 사고 발생 시기 및 장소가 일치하는 터널 상세현황 및 터널소통 정보 (고속도로 공공데이터 포털에서 제공)를 매칭한 뒤, 각 터널 사고 케이스 별로 터널 상세현황 정보로부터 수집한 방재 시설 현황 및 터널 기하구조 데이터와 터널 소통정보로부터 수집한 터널 튜브별 일 교통량 데이터를 보강하였다.

Table 2. Variables used in this study

Variable features		No. of variables in a feature
Dependent	Crash severity	1
Independent	Tunnel disaster prevention facility related	8
	Tunnel geometric characteristic related	7
	Collision related	9
	Traffic related	9
	Weather and temporal factor related	6
	At-fault vehicle and driver related	3

그 결과, 터널 사고 심각도에 영향을 미칠 수 있는 위험요인 후보로 고려된 독립변수의 수가 Table 2에서 나타난 것처럼 상당히 많았다. 연속형 혹은 명목형 변수들의 범주 구분은 터널 위험도 지수 세부평가항목 값들의 범주 구분을 따랐다. 더불어, 수집된 터널 사고데이터에서 전체 사고데이터 케이스 기준 25% 이상의 정보가 누락된 변수들은 위험요인 후보에서 제외하였다.

4. 분석방법

4.1 분석 진행 단계

Table 2에서 볼 수 있는 바와 같이 터널 위험요인으로 변수 42개 모두를 모형에 포함하는 것은 작업의 효율을 떨어뜨리고 변수 간의 다중공선성을 일으켜 결과의 신빙성을 떨어뜨린다. 따라서 본 연구 분석과정의 첫 번째 단계는 많은 독립변수 후보 중 고속도로 터널 사고의 심각도에 상대적으로 큰 영향을 끼치는 주요 위험요인 목록을 먼저 추출하는 것이다. 이를 위하여 본 연구에서는 기계학습 기법의 하나인 RF 알고리즘과 전통적 변수선택 방식인 VIF 기법을 통하여 각각 주요 위험요인 목록을 일차적으로 추출하였다. 두 번째 단계로, RF와 VIF를 이용하여 추출한 각각의 위험요인 목록 전체를 BLR 모형에 포함하되 단계적 선택법 (Stepwise Selection)을 이용하여 각각의 최종 다항 BLR을 구축하였다. 마지막 단계로, 두 번째 단계에서 구축한 각각의 최종 RF 기반 다항 BLR과 VIF 기반 다항 BLR을 전체 모델 적합도, 예측의 정확성 및 독립변수의 유의미성을 기준으로 비교하여 더 우수한 최적 모델을 선정, 선정된 최적 모델을 기반으로 사고 심각도 확률을 예측하였는데 관련된 이론적 원리는 다음과 같다.

4.2 주요 위험요인 추출: RF vs. VIF

RF 알고리즘은 주로 회귀 혹은 분류분석에 이용되는 앙상블 학습 방법의 일종으로 DT를 학습시킬 때 변수선택과 데이터 선택의 임의성을 통하여 DT의 불안정성을 개선한 방법론이다. 기존의 DT 학습데이터에 따라 결과가 생성 트리가 매우 달라지는 불안정성을 극복하기 위한 기법인 배깅(Bagging)은 복원추출에 의한 훈련표본을 만들고 부스트래핑(Bootstrapping) 각 훈련표본에 모델을 적용한 결괏값의 평균 혹은 다수결로 투표하여 최종결과를 도출함으로써 분산을 줄일 수 있는 기법이

다. RF 알고리즘은 트리의 노드 분할 시 무작위로 선택한 일부 설명변수만을 사용하여 (배깅은 모든 설명변수를 사용) 트리 간의 상관성을 낮춤으로써 예측 오차의 분산을 낮춘다. RF의 성능은 트리를 훈련시킨 후 붓스트랩한 데이터들과 원본 데이터를 비교하여 특정 데이터 케이스를 포함하지 않는 붓스트랩 데이터를 따로 모아 RF로 학습을 시켜 나온 결과들의 합인 Out-of-Bag 에러 (이하 OOB)의 수치로 파악을 한다[22]. RF는 변수의 데이터 형태와 변수 간의 다중공선성 이슈에 덜 민감할 뿐만 아니라 변수의 중요성 선정이 가능하다. 변수의 중요성은 해당 변수를 트리 분할에 사용함으로써 얼마나 오류율 혹은 불순도가 낮아졌는가를 기준으로 평가하는데, RF 알고리즘에서는 주로 평균 오류율 감소 (Mean Decrease in Accuracy, 이하 MDA) 혹은 평균 불순도 감소 (Mean Decrease in Gini, 이하 MDG)를 변수의 중요성을 평가하기 위해 사용한다[23].

본 연구의 목적과 데이터 상황을 고려하여 많은 독립 변수 중에서 종속변수인 터널 사고 심각도에 상대적으로 큰 연관이 있는 주요 위험요인들을 선택하는데 RF 알고리즘을 활용하였다. 한편, 다중공선성을 해결하기 위한 전통적인 척도로서 다항 회귀모형에서 VIF를 주로 이용해왔다. VIF는 1-결정계수(R^2)의 역수로 계산되며 일반적으로 그 값이 10이 넘으면 독립변수 간에 다중공선성이 있다고 판단하고 5가 넘으면 주의할 필요가 있다고 본다[24]. 본 연구에서는 적극적으로 다중공선성 이슈를 차단하기 위해 VIF의 최댓값을 5로 선정하여 VIF 값이 5를 넘는 변수들은 모형에서 제외하였다.

4.3 BLR

본 연구에서는 터널 사고 심각도와 관련된 위험요인 목록 추출을 위하여 RF 혹은 VIF를 이용하였으나 이렇게 추출된 위험요인들이 터널 사고 심각도에 미치는 영향을 명확한 수치로 표현할 수 없어 BLR을 추가로 실행하여 터널 사고 심각도 확률예측 모델을 구축하였고 일반적인 BLR 식은 Eq. (1)과 같다.

$$P(Y=1 | X) = \frac{e^{\beta_0 + \beta X}}{1 + e^{\beta_0 + \beta X}} \quad (1)$$

Where, $P(Y=1)$ = probability of injury crash occurrence, β_0 = constant, β = a set of coefficients for independent variables, X = a set of independent variables.

5. 분석결과

5.1 주요 위험요인 목록 추출

트리 500개가 생성될 때까지 RF를 실행하였으며 생성된 트리의 개수가 100개가 된 이후부터 OOB 오류율이 안정되어 21.2%의 OOB 오류율을 보였다. 본 연구와 비슷한 변수의 종류와 데이터 형태를 가지고 RF를 이용, 주요 변수를 선정한 선행연구 논문에서 변수들의 중요도 순위화를 위하여 MDG 대신 MDA를 활용하였는데[4], 본 연구에서도 MDA 기준으로 선택된 변수들이 다양하여 여러 터널 위험요인들의 영향을 확인할 수 있었기 때문에 MDA에 의하여 주요 위험요인들을 순위화하였고 그 결과는 Fig. 3과 같다.

Fig. 3에 의하면 MDA의 값이 5 부근 (붉은 점선)에서 상대적으로 하락추세가 큰 변화 없이 안정되는 것을 볼 수 있었다. 따라서, MDA 값이 5 이상이 되는 상위 18개의 변수가 총 42개의 변수 후보 중 터널 사고 심각도에 상대적으로 큰 영향을 미치는 주요 위험요인들이라 볼 수 있다. MDA 기준에 의하여 선택된 상위 18개의 잠재적 주요 위험요인들의 목록은 다음과 같다: 1) 방재시설 관련: 비상 주차대 및 비상 전화/방송시스템/CCTV 설치 여부; 2) 터널특성 관련: 폭, 높이, 차선 수, 연장, 유사시 보행 가능 폭; 3) 사고환경 관련: 충돌형태, 주 사고원인, 터널 입/출구 펜스, 커브 지역, 도로가 절개지, 터널 존, 야간사고, 터널 주변 공사 여부; 4) 교통흐름 관련: 주행거리계, 교통류 속도; 5) 차량 관련: 가해 차종.

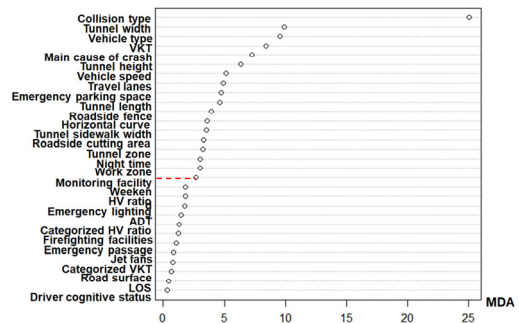


Fig. 3. Variable importance by MDA in RF

반면, VIF를 이용한 주요 위험요인 목록 추출 결과, VIF 값이 5를 넘는 변수들이 다음의 6개로 나타났다: 범주화된 터널 연장, 터널 높이, 범주화된 주행거리계, 튜브당 일 교통량, 범주화된 LOS, 범주화된 대형차 혼입률

(Heavy Traffic Ratio). 이들 변수는 VIF에 의한 위험요인 목록선정 기반 BLR 모형에 독립변수로 사용될 위험요인 목록으로부터 제외되었다.

5.2 주요 위험요인 목록기반 BLR

RF와 VIF에 의하여 각각 도출된 터널사고 심각도에 영향을 주는 위험요인 목록을 독립변수로 이용하여 개별 BLR 모형을 구축한 결과는 Table 3과 같다. Table 3에서 두 모형 모두 데이터에 대한 적합성, 결정계수(Pseudo R2)에서 거의 차이가 없었고 예측의 정확성에서도 약 85%로 비슷함을 보였다. 그러나, RF 기반 BLR 모형은 심각도가 더 높은 부상 사고 발생확률 예측 (Sensitivity=19%)에서 VIF 기반 BLR 모형 (Sensitivity=15%)보다 높은 정확성을 나타내는 것으로 나타났다. 따라서 RF 기반 BLR 모형이 독립변수의 범위를 일차적으로 줄였음에도 독립변수 모두를 쓴 VIF 기반 BLR 모형과 비교하여 예측의

Table 3. RF-based BLR model vs. VIF-based BLR model

Variable (Base)	Variable category	RF-based		VIF-based	
		β	p	β	p
Constant	N/A	-3.97	0.01	-5.11	0.01
Collision type (Single vehicle involved)	Head-on	1.54	0.01	1.73	<0.01
	Angle	1.87	<0.01	1.82	<0.01
	Rear-end	1.89	<0.01	2.08	<0.01
	Facility collided	1.36	<0.01	1.39	<0.01
Tunnel width	Tunnel width	0.25	<0.01	0.25	<0.01
Vehicle type (Car)	Truck	0.79	<0.01	0.84	<0.01
Main cause of crash (Other)	Driver's sleep	2.12	0.05		
Tunnel height, h	Tunnel height	-0.43	0.02		
Tunnel zone (Inner)	Exit	0.52	0.05		
Emergency facilities (None)	Call/Broadcast/CCTV	-0.99	<0.01	-1.19	<0.01
Road surface (Dry)	Wet			-1.19	0.03
Driver status (Sobriety)	Fatigue			1.04	<0.01
Categorized tunnel height (h < 5m)	h ≥ 7.5m			-1.16	0.02

Pseudo R² (RF vs. VIF) = 0.27 vs. 0.26, Sensitivity (RF vs. VIF) = 19.0% vs. 15.1%, Specificity (RF vs. VIF) = 97.4% vs. 98%, Total accuracy (RF vs. VIF) = 85.0% vs. 84.7%.

정확성에서 약간의 우위를 보이는 결과를 보임에 따라 RF 알고리즘으로 일차 주요 독립변수만을 추출 후 이 목록만을 사용하여 시행한 BLR은 예측성능이 충분한 효율적 모형화 과정임이 확인되었다.

따라서 본 연구의 목적과 모델링의 효율성, 모델 전체의 적합성, 예측의 정확성을 통합적으로 고려하여 RF 기반 BLR 모형을 더 우수한 모델로 선정하였다. 최적 모델인 RF 기반 BLR 모형에 포함된 유의미한 위험요인들에 초점을 두고 터널 부상 사고에 미치는 각 위험요인의 구체적인 영향을 다음과 같이 살펴보았다.

RF 기반 BLR 모형에서 사고환경 및 차량 관련 위험요인으로 2개 이상 차량의 정면충돌/측면충돌/추돌/도로시설(중분대/도로가 펜스) 충돌, 졸음운전, 터널 출구, 트럭은 터널에서의 부상 사고 발생확률을 높이는 것으로 나타났다. 또한, 터널구조 관련 위험요인으로 터널의 폭이 넓을수록, 터널의 높이가 낮을수록 부상 사고 발생확률이 높아지는 것으로 나타났다. 특히 터널의 폭은 주행거리계를 구성하는 터널의 일 교통량과 양의 상관관계가 있는 것으로 확인되어 터널의 폭은 주행거리계를 대신할 수 있는 변수임이 확인된다. 긴급전화, 비상방송시설 혹은 CCTV 설치 여부는 터널 방재 시설 중 유일하게 RF 기반 BLR 모형에 포함된 유의미한 요인으로, 이러한 방재 시설이 설치된 터널에서는 그렇지 않은 터널에 비하여 부상 사고의 확률이 줄어드는 것으로 나타났다.

5.3 터널 안전관리 정책지원에 대한 고찰

3.2 에서 서술한 기존의 터널 위험도 지수 세부평가항목 리스트와 5.1에서 RF 알고리즘에 의해 일차적으로 도출된 주요 위험요인 목록을 비교해 보았을 때, 기존의 세부평가항목 중 주행거리계, 터널 높이, 곡선반경이 공통적인 것으로 나타났다. 이에 주행거리계, 터널 높이, 곡선반경은 터널 사고의 위험요인으로서 타당함이 확인되었다. 또한, 사고 가해 차량이 트럭이나 트레일러 형태의 대형차량일 경우 부상 사고 발생확률을 높이는 데 유의미한 것으로 나타났는데 이는 기존의 대형차 관련 항목들도 터널 위험도 지수 세부평가항목으로서 타당함을 암시한다.

그러나, 기존의 터널 위험도 지수 세부평가항목 중 터널특성 및 정체 정도와 관련된 항목으로 포함된 진입부경사도 및 LOS 항목은 5.1에서 보듯이 RF 알고리즘 기반 주요 위험요인 추출에서 변수로 선정되지 않은 대신 다음의 변수들이 포함되었다: 대피, 통신, 모니터링 관련 방재 시설, 터널의 폭/차로 수/연장, 충돌 전 차속, 야간

사고, 터널 주변 도로 펜스 유무, 절개지 유무, 공사 여부. 이러한 결과는 터널 위험도 지수 세부평가항목으로 대피, 통신, 모니터링, 조명 관련 방재 시설, 터널 폭, 교통류 속도, 도로가 시설, 산사태 가능성 및 공사 관련 변수를 터널 위험도 지수 세부평가항목으로 추가 검토할 필요가 있음을 시사한다.

또한, 실제 부상 사고가 일어난 고속도로 전체 터널 중 RF 기반 BLR 모형에 의해 부상 사고 발생확률이 0.5 보다 높게 예측된 터널들을 추출한 후, 그중에서 긴급전화, 비상방송설비, CCTV가 모두 설치되지 않은 터널을 확인한 결과 Table 4와 같이 전국 13개의 고속도로 루트에서 총 27개소의 터널이 확인되었다.

Table 4. Freeway tunnels that need to install emergency call/broadcast/CCTV

Route No.	Tunnel name (Bound for)	Route No.	Tunnel name (Bound for)
1	Daejeon (Seoul)	45	Goryeong 3 (Yangpyeong)
10	Joseong 1 (Youngam)		Goryeong 4 (Yangpyeong)
12	Namwon (Gwangju)		Gimcheon 1 (Yangpyeong)
15	Mongtan 3 (Mokpo)		Dujeong (Yangpyeong)
	Seulchi (Suncheon)		Mungyeong 2 (Changwon)
27	Shinri 2 (Suncheon)	50	Jungwon (Yangpyeong)
	Osu 1 (Suncheon)		Daegwanryeong 4 (Gangneung)
	Yongam 3 (Wanju)		Daegwanryeong 4 (Incheon)
	Yongam 4 (Wanju)		Masung (Gangneung)
30	Naeseo 4 (Cheongju)	55	Gunwui (Chuncheon)
	Suriti (Sangju)		Iljik (Chuncheon)
40	Banje (Pyeongtaek)	60	Misa (Yangyang)
			Songsan (Seoul)
		100	Sorae (Pangyo)
		110	Gwangmyeong (Incheon)

더불어 RF 기반 BLR 모형을 이용하여 부상 사고 발생확률을 예측한 결과 긴급전화, 비상 방송설비, CCTV 설치가 부상 사고 발생확률을 낮추는데 유의미하게 영향

을 준다는 것이 확인되었다. 이러한 결과는 타 방재 시설 대비 통신 및 모니터링 설비를 먼저 보강하는 것이 터널 사고의 심각성을 완화하는데 효율적임을 암시한다.

6. 결론 및 후속연구

본문의 서론에서 제시된 목적을 달성하기 위해 본 연구는 2012년부터 2016년까지 국내 전체 고속도로 터널에서 발생한 교통사고 데이터에서 누락된 터널의 구조적 특성, 교통류 속도, 터널 상세현황을 추가 수집, 사고와 매칭하여 확장된 고속도로 터널 사고데이터 세트를 구축하였다. 이어 확장된 터널 사고데이터의 많은 변수 중에서 터널 사고 심각도에 상대적으로 큰 영향을 미치는 주요 위험요인을 RF 알고리즘을 적용하여 추출, 기존의 터널 위험도 지수 세부평가항목 리스트와 비교하였다. 더불어 방재 시설을 먼저 보강할 터널과 해당 방재 시설을 선정하기 위해 RF와 VIF 기반 BLR 모형을 비교하여 최적의 BLR 모형을 선택한 후, 터널 부상 사고 발생확률을 예측하였고 다음의 결과를 얻었다.

첫째, 기존 위험도 지수 세부평가항목으로서 타당성이 확인된 위험요인은 주행거리계, 터널 높이, 곡선반경, 대형차 관련 세부평가항목으로 나타났다. 또한, RF 기반 BLR 모델은 변수선택의 효율성 및 터널 부상 사고 예측의 정확성 면에서 우수한 모델임이 확인되었다. 둘째, 긴급전화, 비상방송시설 및 CCTV와 같은 통신 및 모니터링 관련 시설들이 다른 방재 시설보다 터널 사고 심각도에 상대적으로 큰 영향을 미치는 것으로 확인되었다. 셋째, 총 13개 고속도로 루트의 27개 터널에 먼저 긴급전화, 비상방송시설 및 CCTV와 같은 통신 및 모니터링 관련 방재 시설 설치가 필요함이 확인되었다.

본 연구에서는 탑승자의 안전띠 여부와 운전자 나이 정보가 전체 터널 사고 케이스의 25% 이상이 누락되어 있어 모형화 과정 중 위험요인 후보에서 제외되었고, 입/출구 표고 차와 같은 터널특성 및 방재 시설 일부의 정보수집에 제한이 있었다. 또한, 본 연구에서는 고속도로 터널에서만 발생한 사고를 중심으로 분석하였다. 따라서 터널 위험도 등급 재산정을 위한 범용성 있는 터널 위험도 지수 세부평가항목 리스트 확정을 위하여 후속연구에서는 터널에서 발생한 사고와 방재 시설 상세설치 현황과 관련하여 데이터를 추가로 수집해야 할 뿐만 아니라, 국내 모든 도로 터널에서 발생한 교통사고의 빈도와 심각도를 결합하여 터널 위험요인을 추출할 필요가 있을 것이다.

References

- [1] Korea Ministry of Land, Infrastructure and Transport (KMOLIT). Road Transportation Safety Enhancement Strategies with Road Infrastructure Improvement [Internet], KMOLIT, 2020, Available From: http://www.molit.go.kr/srocm/USR/N0201/m_13551/dtl.jsp?id=95083853 (accessed Jan. 30, 2021)
- [2] KMOLIT. Guideline for Road Tunnel Disaster Prevention Facility Installation and Management [Internet], 2009, Available From: <https://www.law.go.kr/> (accessed Feb. 11, 2021)
- [3] KMOLIT, Yearbook of Road Bridge and Tunnel Statistics 2020, Report of Statistics, KMOLIT, Korea, pp. 48.
- [4] X. Wang, S. Kim, "Prediction and factor identification for crash severity: comparison of discrete choice and tree-based models", *Transportation Research Record*, Vol.2673, No.9, pp.640-652, 2019.
DOI: <https://doi.org/10.1177/0361198119844456>
- [5] M. Chen, M. Chen, "Modeling road accident severity with comparisons of logistic regression, decision tree and random forest", *Information*, Vol.11, No.5, 270, 2020.
DOI: <https://doi.org/10.3390/info11050270>
- [6] G. Pillajo-Quijia, B. Arenas-Ramírez, C. González-Fernández, F. Aparicio-Izquierdo, "Influential factors on injury severity for drivers of light trucks and vans with machine learning methods", *Sustainability*, Vol.12, No.4, 1324, 2020.
DOI: <https://doi.org/10.3390/su12041324>
- [7] I. Cho, D. Han, S. Kim, J. Yoon, "A Study on the methodology modelling of risk assessment in road tunnels", *Journal of the Korea Academia-Industrial Cooperation Society*, Vol.17, No.4, pp.59-73, 2016.
DOI: <http://dx.doi.org/10.5762/KAIS.2016.17.4.59>
- [8] H. Lee, D. Joo, C. Hyun, J. Jeong, B. Park et al., "Study on the analysis for the effects of the section speed enforcement system at the Misiryong tunnel section", *J. Korea Inst. Intell. Transp. Syst.*, Vol.12, No.3, pp.11-18, 2013.
DOI: <http://dx.doi.org/10.12815/kits.2013.12.3.011>
- [9] S. Lee, C. Lee, "Development of VSL control algorithms for various traffic conditions in tunnels", *J. Korea Inst. Intell. Transp. Syst.*, Vol.18, No.3, pp.84-94, 2019.
DOI: <https://doi.org/10.12815/kits.2019.18.3.84>
- [10] D. Lee, J. Kim, Y. Jang, J. Hwang, M. Kim, "Lessons learned from serious tunnel traffic accidents in winter", *Magazine of Korean Tunnelling and Underground Space Association*, Vol.22, No.2, pp.55-62, 2020.
- [11] C. Baek, S. Park, B. Choi, "Analysis of the relationship between human risk factors and evacuation behavior for tunnel safety", *Journal of the Korea Academia-Industrial Cooperation Society*, Vol.20, No.7, pp.535-546, 2019.
DOI: <https://doi.org/10.5762/KAIS.2019.20.7.535>
- [12] H. Huang, Y. Peng, J. Wang, Q. Luo, X. Li, "Interactive risk analysis on crash injury severity at a mountainous freeway with tunnel groups in China", *Accident Analysis and Prevention*, Vol.111, pp.56-62, 2018.
DOI: <https://doi.org/10.1016/j.aap.2017.11.024>
- [13] J. Park, Y. Yoo, B. Park, "An analysis study for reasonable installation of tunnel fire safety facility", *Journal of Korean Tunnelling and Underground Space Association*, Vol.17, No.3, pp.243-248, 2015.
DOI: <http://dx.doi.org/10.9711/KTAJ.2015.17.3.243>
- [14] J. Kim, S. Min, "Study on securing safety based on emergency power dualization inside road tunnel", *J. Korean Soc. Hazard Mitig.*, Vol.20, No.1, pp.217-221, 2020.
DOI: <https://doi.org/10.9798/KOSHAM.2020.20.1.217>
- [15] K. Kim, M. Cho, T. Kim, "Effect of re-ventilated fan capacity on road tunnel fire", *Journal of the Korea Academia-Industrial Cooperation Society*, Vol.20, No.9, pp.204-210, 2019.
DOI: <https://doi.org/10.5762/KAIS.2019.20.9.204>
- [16] J. Yoo, D. Rie, H. Shin, "A study of classification of road tunnel for fire safety", *Journal of the Korean Society of Safety*, Vol.20, No.3, pp.112-119, 2005.
- [17] J. Yoo, A Study to Revise Road Tunnel Disaster Facility Installation and Management Guideline: Final Report, Technical Report, KMOLIT, Korea, Report. No. TRKO201600016293, 2015.
- [18] J. Wang, A. Pervez, Z. Wang, C. Han, L. Hua et al., "Crash analysis of Chinese freeway tunnel groups using a five-zone analytic approach", *Tunnelling and Underground Space Technology*, Vol.82, pp.358-365, 2018.
DOI: <https://doi.org/10.1016/j.tust.2018.08.037>
- [19] S. Bassan, "Overview of traffic safety aspects and design in road tunnels", *IATSS Research*, Vol.40, No.1, pp.35-46, 2016.
DOI: <https://doi.org/10.1016/j.iatssr.2016.02.002>
- [20] A. Pervez, H. Huang, C. Han, J. Wang, Y. Lia, "Revisiting freeway single tunnel crash characteristics analysis: a six-zone analytic approach", *Accident Analysis and Prevention*, Vol.142, 105542, 2020.
DOI: <https://doi.org/10.1016/j.aap.2020.105542>
- [21] Department of Advanced Road Safety, Tunnel Safety Management Enhancement Strategies [Internet], KMOLIT, 2016, Available From: http://www.molit.go.kr/USR/NEWS/m_71/dtl.jsp?id=95078508 (accessed Jan. 30, 2021)
- [22] L. Breiman, "Random forest", *Machine Learning*, Vol.45, pp.5-32, 2001.
DOI: <https://doi.org/10.1023/A:1010933404324>
- [23] X. Qiu, P. Choi, "A study on discrimination in mortgage lending in the United States: a revisit by random forest method", *Journal of the Korean Data*

and Information Science Society, Vol.30, No.2,
pp.261-270, 2019.

DOI: <http://dx.doi.org/10.7465/jkdi.2019.30.2.261>

- [24] J. Hair, R. Anderson, R. Tatham, W. Black,
Multivariate data analysis, 3rd ed., Pearson Prentice
Hall Publisher, 1992.
-

정 소 영(Soyoung Jung)

[정회원]



- 2004년 3월 : The Ohio State University 도시및지역계획학과 (도시및지역계획학석사)
- 2010년 8월 : University of Wisconsin-Madison 건설및환경공학과 (교통공학박사)

- 2017년 3월 ~ 현재 : 동양대학교 동두천캠퍼스 스마트안전시스템학부 교수

〈관심분야〉

교통안전, 교통운영, 교통정보