ISSN : 1226-7244 (Print)
ISSN : 2288-243X (Online)
j.inst.Korean.electr.electron.eng.Vol.25,No.3,571~577,September 2021
논문번호 21-03-25
https://doi.org/10.7471/ikeee.2021.25.3.571
160

자율주행자동차의 안전 및 보안을 위한 설계 및 검증 표준: ISO/TR 4804

Design and Verification Standard for Safety and Cybersecurity of Autonomous Cars: ISO/TR 4804

이 성 수*★

Seongsoo Lee*★

Abstract

This paper describes ISO/TR 4804, an international standard to describe how to design and verify autonomous cars to ensure safety and cybersecurity. Goals of ISO/TR 4804 are (1) positive risk balance and (2) avoidance of unreasonable risk. It also 12 principles of safety and cybersecurity to achieve these goals. In the design procedures, it describes (1) 13 capabilities to achieve these safety and cybersecurity principles, (2) hardware and software elements to achieve these capabilities, and (3) a generic logical architecture to combine these elements. In the verification procedures, it describes (1) 5 challenges to ensure safety and cybersecurity, (2) test goals, platforms, and solutions to achieve these challenges, (3) simulation and field operation methods, and (4) verification methods for hardware and software elements. Especially, it regards deep neural network as a software component and it describe design and verification methods of autonomous cars.

요 약

본 논문에서는 자율주행자동차의 안전성 및 보안성을 보장하기 위해서 설계하고 검증하는 방법을 규정한 국제 표준인 ISO/TR 4804에 대해 다룬다. ISO/TR 4804는 자율주행자동차가 (1) 인간 운전자보다 훨씬 더 안전하고 (2) 타당하지 않은 위험이 없도록 하는 것을 목표로 하며, 이를 위해 12개의 안전성 및 보안성 원칙을 제시한다. 설계 과정에서는 (1) 안전성 및 보안성 원칙을 달성하는데 필요한 13개의 역량, (2) 이 역량을 수행하기 위해 필요한 하드웨어 및 소프트웨어 요소, (3) 이 요소를 결합한 논리적, 일반적인 아키텍처 등을 규정한다. 검증 과정에서는 (1) 안전성 및 보안성을 검증하기 위한 5개의 과업, (2) 이 과업을 완수하기 위한 테스트 목표, 플랫폼, 솔루션, (3) 시뮬레이션 방법 및 필드 운영 방법, (4) 하드웨어 및 소프트웨어 요소의 검증 방법 등을 규정한다. 특히 심층 신경망을 하나의 소프트웨어 요소로 간주하고, 심층 신경망이 적용된 자율주행자동차를 설계하고 검증하는 방법을 규정한다.

Key words : ISO/TR 4804, Safety, Security, Autonomous Driving, ISO 26262, ISO/PAS 21448, ISO 21434

---

* Soongsil University (Professor)
★ Corresponding author
E-mail : sslee@ssu.ac.kr, Tel : +82-2-820-0692

※ Acknowledgment
This work was supported by Korea Institute for Advancement of Technology (KIAT) grant funded by the Korea Government (MOTIE). (P0017011, HRD Program for Industrial Innovation) This work was also supported by Industrial Technology Challenge Track (20012624) of the Ministry of Trade, Industry and Energy (MOTIE) / Korea Evaluation Instutite of Industrial Technology (KEIT).
Manuscript received Sep. 21, 2021; revised Sep. 24, 2021; accepted Sep. 27, 2021.

Ⅰ. 서론

인간 운전자 대신에 인공지능이 운전을 수행하는 자율주행자동차의 등장에 따라 어떻게 자율주행자동차의 안전성과 보안성을 보장할 수 있는지가 매우 중요해지고 있다. 자율주행자동차에 사용되는 인공지능은 복잡성이 매우 높으며 동작을 분석하기 어렵고 예측하지 못한 위험 요소가 많고 설계 및 테스트가 규격화기 어렵다. 자율주행자동차의 안전성과 보안성을 보장하지 못하면 해당 제품의 상용화에 큰 지장을 초래할 뿐만 아니라 경우에 따라 제조사가 심각한 법적 책임을 감수해야 할 수 있다[1][2].

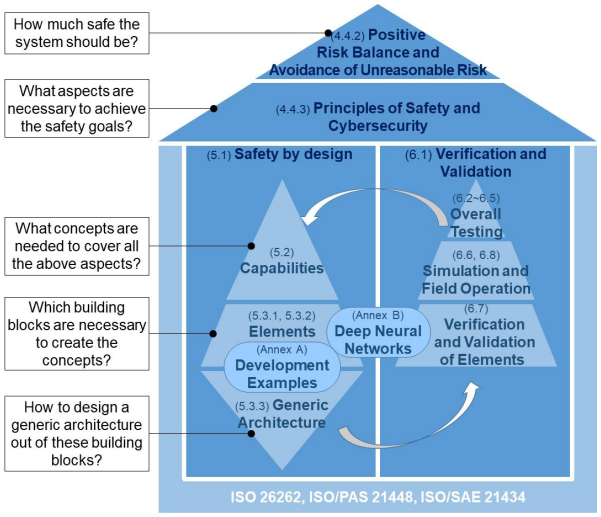

Fig. 1. Overview of ISO/TR 4804 [7].
그림 1. ISO/TR 4804의 개요 [7]

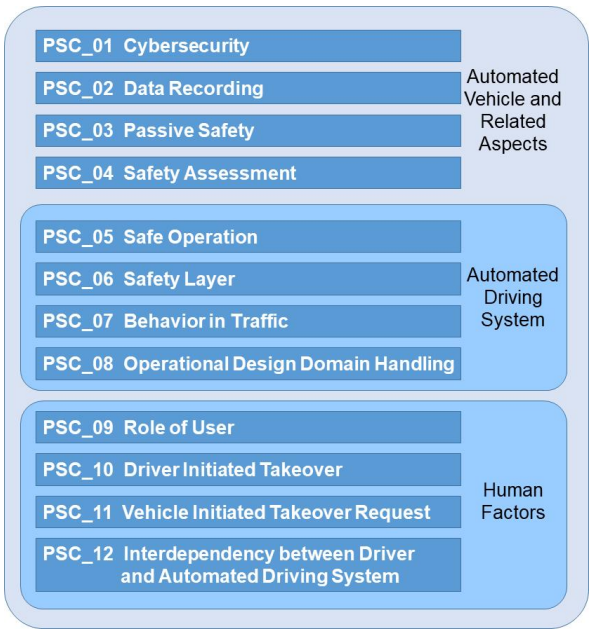

Fig. 2. Principles of safety and cybersecurity [7].
그림 2. 안전성 및 보안성 원칙 [7]

자동차에서 안전성과 보안성을 보장하기 위한 국제 표준으로는 ISO 26262[3], ISO/PAS 21448[4], ISO/SAE 21434[5] 등이 있으나 자율주행과 인공지능에 대한 고려가 크게 부족한 편이다. 최근 UL (Underwriters Laboratories)에서 자율주행자동차의 안전성을 평가하기 위한 국제 표준인 UL 4600[6]을 제정하였으나 이 표준은 주로 안전성과 보안성을 증명하는 방법에 초점이 맞춰져 있으며 설계와 검증에 대해서는 잘 나타나 있지 않다. 이에 따라 국제표준화 기구(ISO: International Standard Organization)에서는 자율주행자동차의 안전성 및 보안성을 보장하기 위해서 설계하고 검증하는 방법을 규정한 ISO/TR 4804[7]을 제정하였다. 본 논문에서는 ISO/TR 4804에서 안전성과 보안성을 보장하기 위해서 설계하고 검증하는 방법을 자세히 살펴본다.

Ⅱ. ISO/TR 4804의 개념 및 구조

ISO/TR 4804는 벤츠, 아우디, 폭스바겐, BMW 등 자동차 제조사와 인피니언, 컨티넨털, 인텔 등 자동차 부품사가 2019년 공동으로 집필한 백서(white paper)[8]로 시작하여 2020년에 ISO에서 기술 보고서(TR: technical report)로 승인되었으며 현재 국제 표준(IS: international standard)으로 제정 작업이 진행 중이다. 이 표준은 SAE J3016[9]에

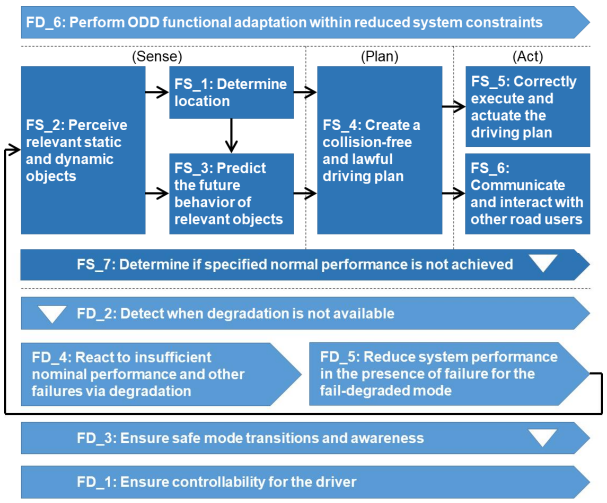

Fig. 3. Fail-safe capabilities and fail-degraded capabilities [7].
그림 3. 고장 안전 역량 및 고장-저하 역량 [7]

ID	PSC_01 Cybersecurity	PSC_02 Data recording	PCS_03 Passive safety	PSC_04 Safety assessment	PSC_05 Safe operation	PSC_06 Safety layer	PSC_07 Behaviour in traffic	PSC_08 Operational design domain handling	PSC_09 Role of user	PSC_10 Driver initiated takeover	PSC_11 Vehicle initiated takeover request	PSC_12 Interdependency between driver and ADS
FS_1 Determine location	×			×			×	×				
FS_2 Perceive relevant static and dynamic objects	×			×			×					
FS_3 Predict the future behaviour of relevant objects	×			×			×					
FS_4 Create a collision-free and lawful driving plan	×			×			×					
FS_5 Correctly execute and actuate the driving plan	×			×			×					
FS_6 Communicate and interact with other road users	×			×			×					
FS_7 Determine if specified nominal performance is not achieved	×			×		×		×				
FD_1 Ensure controllability for the driver	×			×	×				×	×	×	×
FD_2 Detect when degradation is not available	×			×	×							
FD_3 Ensure safe mode transitions and operating mode awareness	×			×	×	×			×	×	×	×
FD_4 React to insufficient nominal performance and other failures via degradation	×			×	×	×						
FD_5 Reduce system performance in the presence of failure for the fail-degraded mode	×			×	×	×						
FD_6 Perform ODD functional adaptation within reduced system constraints	×			×	×	×		×			×	

Fig. 4. Fail-safe and fail-degraded capabilities to satisfy safety and cybersecury principles [7]..
그림 4. 안전성 및 보안성 원칙을 만족하는데 필요한 고장-안전 역량 및 고장-저하 역량 [7]

서 정의한 자율주행의 3단계 이상, 즉 인공지능이 인간 운전자 대신에 운행의 전부 또는 대부분을 담당하는 경우에 적용된다.

ISO/TR 4804은 그림 1과 같이 자율주행자동차가 (1) 인간 운전자보다 훨씬 더 안전하고(positive risk balance) (2) 타당하지 않은 위험이 없도록(avoidance of unreasonable risk) 하는 것을 안전 목표로 하며, 이를 위해 안전성 및 보안성 원칙(principles)을 제시한다. 설계 과정에서는 (1) 안전성 및 보안성 원칙을 달성하는데 필요한 역량(capabilities), (2) 이 역량을 수행하기 위해 필요한 하드웨어 및 소프트웨어 요소(elements), (3) 이 요소를 결합한 일반적인 아키텍쳐(generic architecture) 등을 규정한다. 검증 과정에서는 (1) 안전성 및 보안성을 검증하기 위한 과업(challenges)과 이 과업을 완수하기 위한 테스트 솔루션, (2) 시뮬레이션 방법 및 필드 운영 방법, (3) 하드웨어 및 소프트웨어 요소의 검증 방법 등을 규정한다. 또한 다양한 개발 예제와 함께 심층 신경망(DNN: deep neural network)의 사용도 다루고 있다.

안정성 및 보안성 원칙은 그림 2와 같으며, 12개의 원칙은 크게 자율주행 전반에 걸친 이슈, 자율주행시스템에 관련된 이슈, 운전자에 관련된 이슈로 나눌 수 있다. 이러한 원칙을 달성하기 위해 ISO/TR 4804에서는 그림 3과 같은 13개의 역량을 규정하였는데 이들 역량은 크게 고장-안전 역량(fail-safe), 즉 고장나도 최소 안전 조건을 만족하는 역량과 고장-저하(fail-degraded) 역량, 즉 고장나면 성능이 저하되지만 정해진 시간 안에 최소 안전 조건을 만족하는 역량으로 나뉜다. 12개 원칙과 13개 역량은 그림 4와 같이 연관되는데 이를 활용하면 각 원칙을 달성하는데 어떤 역량이 필요한 지

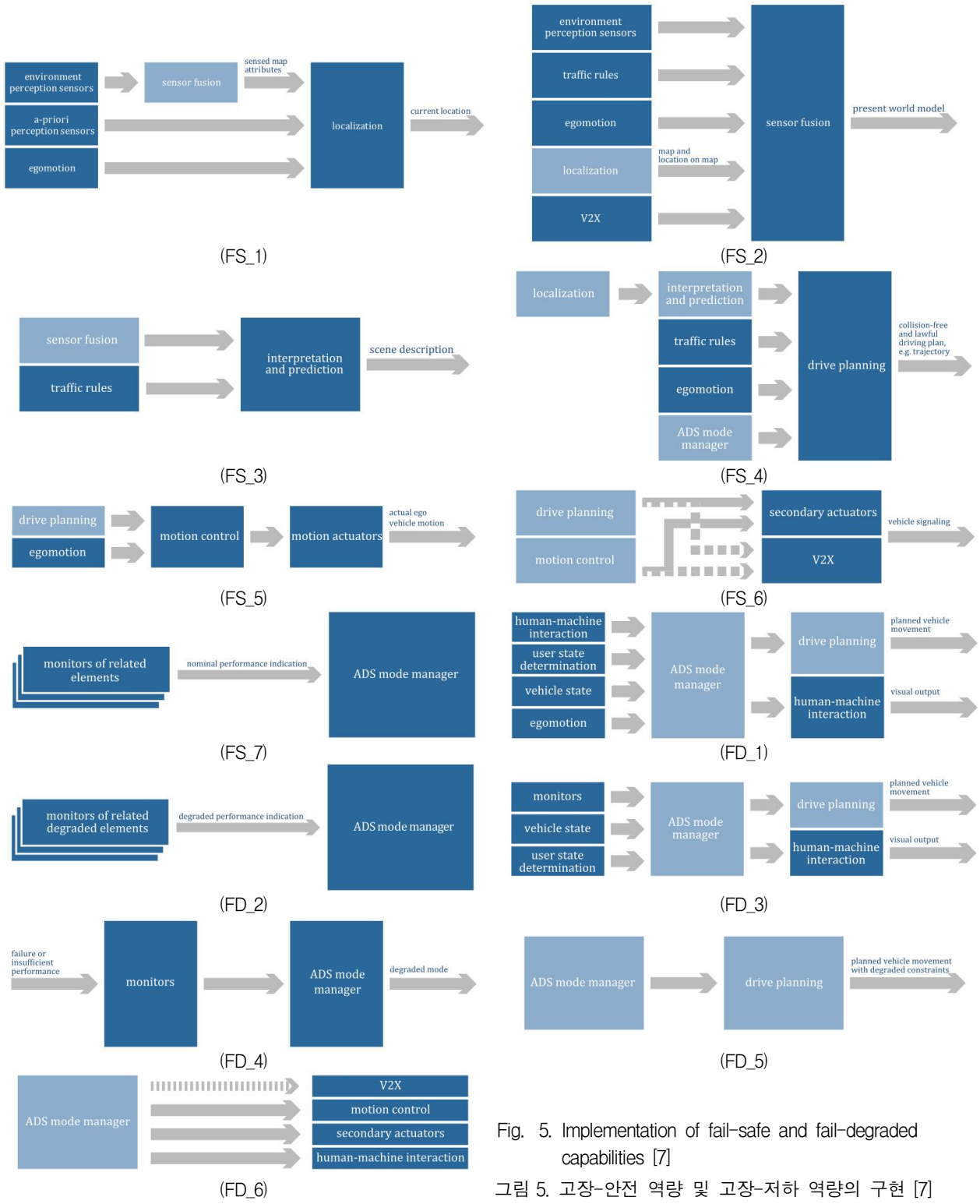

Fig. 5. Implementation of fail-safe and fail-degraded capabilities [7]

그림 5. 고장-안전 역량 및 고장-저하 역량의 구현 [7]

를 파악하여 설계 및 검증 과정을 계획하고 수행하는데 도움이 된다.

III. ISO/TR 4804의 세부 내용

ISO/TR 4804에서는 13개 역량을 그림 5와 같이 구현하도록 규정하였으며, 이를 수행하기 위해 필요한 하드웨어 및 소프트웨어 요소를 다음과 같이 같이 규정하였다. 이들 요소를 결합한 일반적인 시스템 아키텍쳐는 그림 6과 같다.

- 환경 인식 센서(environment perception sensor)

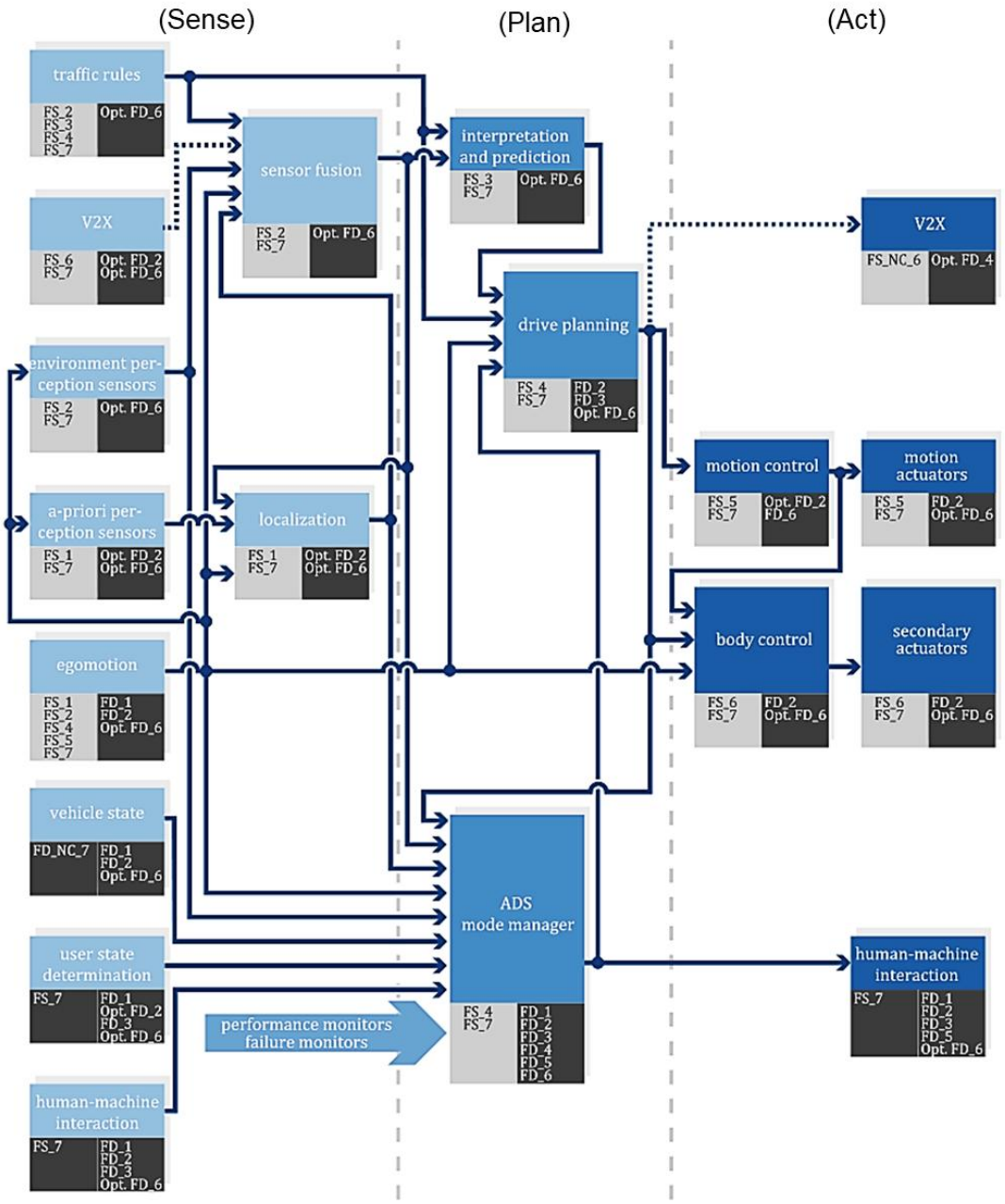

Fig. 6. Generic architecture with hardware and software elements [7].
그림 6. 하드웨어 및 소프트웨어 요소를 결합한 일반적인 아키텍쳐 [7]

- 위치 데이터(a-priori perception sensor)
- 차량 통신(V2X)
- 센서 융합(sensor fusion)
- 해석(interpretation) 및 예측(prediction)
- 위치 측정(localization)
- 자율주행 모드 관리자(ADS mode manager)
- 차량 자체 운동(egomotion)
- 주행 계획(drive planning)
- 교통 규칙(traffic rule)
- 움직임 제어(motion control)
- 움직임 작동기(motion actuator)
- 2차 작동기(secondary actuator)
- 인간-기계 상호작용(HMI)
- 사용자 상태 결정(user state determination)
- 차량 상태(vehicle state)
- 모니터(monitor)
- 연산기(processing unit)
- 전력 공급(power supply)
- 통신 네트워크(communication network)

Test item / Test Platform	Test SW	Test HW	Vehicle	Driver	Driving Environment
Simulation in the closed loop	Virtual	Virtual	Virtual	Virtual	Virtual
	Real			None	
Software reprocessing	Virtual	Virtual	None	None	Virtual
	Real				
Hardware in the closed loop	Real	Real	Virtual	Virtual	Virtual
				None	
Hardware reprocessing	Real	Real	None	None	Virtual
Driver in the loop	Real	Virtual	Virtual	Real	Virtual
			Real		
		None	None		Real
Proving ground	Real	Real	Real	Real	Real
				Robot	
Open road	Real	Real	Real	Real	Real

Fig. 7. Test items and test platforms [7].
그림 7. 테스트 사항과 테스트 플랫폼 [7]

ISO/TR 4804에서는 안전성 및 보안성을 검증하기 위한 과업을 다음과 같이 규정하였다.

- Challenge 1: 운전자의 개입이 없는 상태에서 인간 운전자보다 훨씬 더 안전하고 타당하지 않은 위험이 없을 것에 대한 통계적 증명
- Challenge 2: 운전자와의 인수인계 등 운전자의 개입에서의 시스템 안전
- Challenge 3: 현재까지 알려지지 않은 시나리오에 대한 고려
- Challenge 4: 다양한 시스템 구성 및 변화에 대응한 검증
- Challenge 5: 기계 학습 시스템에 대응한 검증

그림 7은 ISO/TR 4804의 테스트 사항과 테스트 플랫폼을 나타낸 것이며 그림 8은 ISO/TR 4804의 테스트 전략을 나타낸 것이다.

ISO/TR 4804는 심층 신경망을 하나의 소프트웨어 요소로 간주하고, 심층 신경망이 적용된 자율주행자동차를 설계하고 검증한다. 그림 9는 이때의 심층 신경망 모델을 나타난 것이다.

Ⅳ. 결론

본 논문에서는 자율주행자동차의 안전성과 보안성을 보장하기 위해 설계하고 검증하는 방법을 규정한 국제 표준인 ISO/TR 4804에 대해 살펴보았다. ISO/TR 4804는 주요 자동차 제조사와 부품사가 주축이 되어 개발한 국제 표준이며 자율주행자동차에 관련된 다른 국제 표준에 비해 상당히 구체적이고 세부적으로 자율주행자동차의 설계 및 검증 방법을 설명하고 있다. 이 분야는 국내 기술이

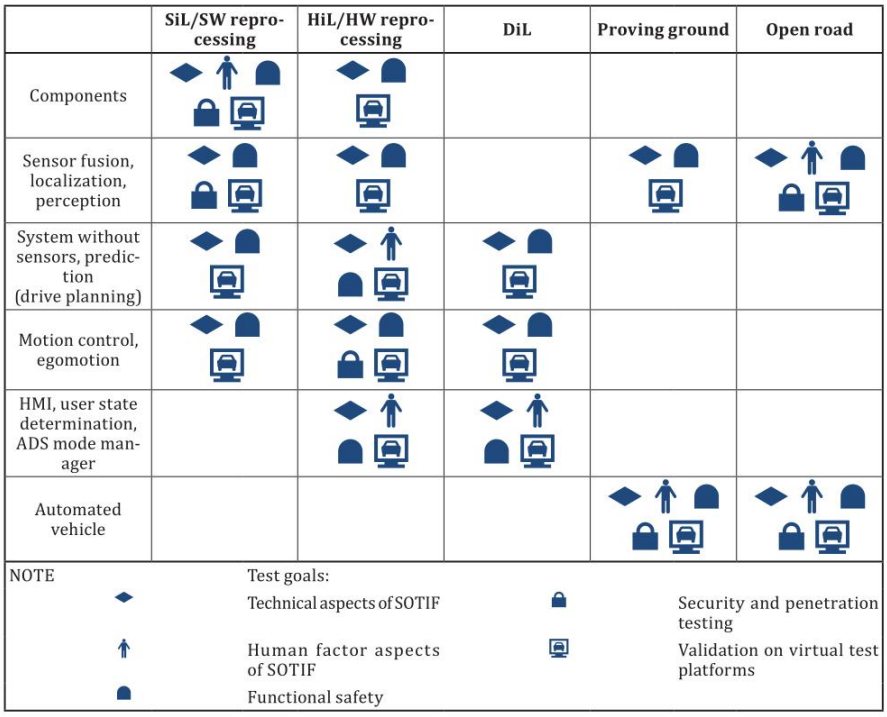

Fig. 8. Test strategies [7].
그림 8. 테스트 전략 [7]

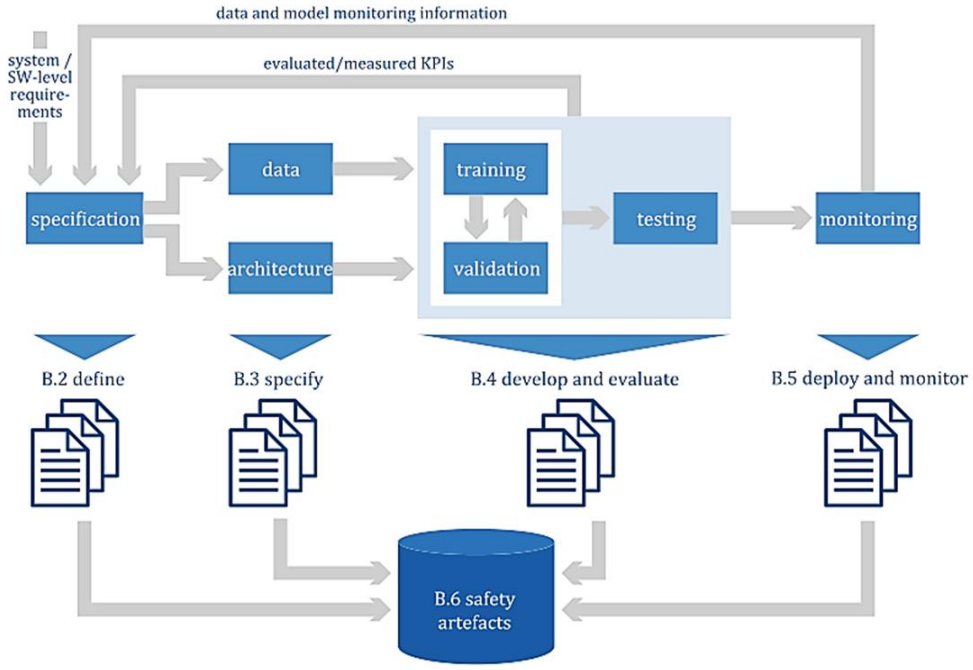

Fig. 9. Deep neural network model [7].
그림 9. 심층신경망 모델 [7]

비교적 취약한 것으로 평가받고 있어서 본 논문을 통해 자율주행자동차 산업의 발달에 기여할 수 있기를 희망한다.

References

[1] S. Ihm and S. Lee, “A Study on the Liability for Damages and the Burden of Proof Relating the Autonomous Car,” Yonsei Law Review, vol.30, no.3, pp.309-332, 2020. DOI: 10.21717/ylr.30.3.10

[2] S. Lee, “ISO 26262 and ISO/PAS 21448 as Exemption Clauses of Product Liability,” *j.inst. Korean.electr.electron.eng.,* vol.23, no.1, pp.346-349, 2019. DOI: 10.7471/ikeee.2019.23.1.346

[3] ISO 26262-1:2018, “Road vehicles - Functional safety - Part 1: Vocabulary,” https://www.iso.org/standard/68383.html

[4] ISO/PAS 21448:2019, “Road vehicles - Safety of the intended functionality,” https://www.iso.org/standard/70939.html

[5] ISO/SAE 21434:2021, “Road vehicles - Cyber security engineering,” https://www.iso.org/standard/70918.html

[6] UL 4600, “Standard for Safety for the Evaluation of Autonomous Products,” https:// ul.org/UL4600

[7] ISO/TR 4804:2020, “Road vehicles - Safety and cybersecurity for automated driving systems - Design, verification, and validation,” https://www.iso.org/standard/80363.html

[8] Aptiv et al., “Safety first for automated driving,” https://www.aptiv.com/docs/default-source/white-papers/safety-first-for-automated-driving-aptiv-white-paper.pdf

[9] SAE J3016_201806, “Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles,” https://www.sae.org/standards/content/j3016_201806