

## 토픽모델링을 이용한 무인지상차량(UGV) 특허 동향 분석

김기환<sup>1)</sup> · 전차수<sup>2)</sup> · 송지훈<sup>3)</sup> · 전정환<sup>\*,2)</sup>

<sup>1)</sup> 국방기술진흥연구소 방위산업전략팀

<sup>2)</sup> 경상국립대학교 산업시스템공학부

<sup>3)</sup> 경상국립대학교 기술경영학과

### Patent Trend Analysis of Unmanned Ground Vehicles(UGV) using Topic Modeling

Kihwan Kim<sup>1)</sup> · Chasoo Jun<sup>2)</sup> · Chiehoon Song<sup>3)</sup> · Jeonghwan Jeon<sup>\*,2)</sup>

<sup>1)</sup> Defense Industry Strategy Team, Korea Research Institute for defense Technology Planning and Advancement, Korea

<sup>2)</sup> Department of Industrial System Engineering, Gyeongsang National University, Korea

<sup>3)</sup> Department of Management of Technology, Gyeongsang National University, Korea

(Received 2 January 2024 / Revised 22 April 2024 / Accepted 27 April 2024)

#### Abstract

This study provides a thorough examination of Unmanned Ground Vehicles(UGVs), focusing on crucial technologies and trends across major global markets. It includes an in-depth patent analysis revealing the dominant positions of the United States and the European Union in this field. Additionally, it underscores substantial advancements made by China, Japan, and Korea since 2010. Using Latent Dirichlet Allocation(LDA)-based patent text mining, the study identified key technology areas in UGV development, such as advanced control systems, navigation technologies, power supply mechanisms, and sensing and communication tools. Through linear regression analysis, the study predicted the future paths of these technology areas, offering important insights into the evolving world of UGV technology. The findings can provide strategic guidance for stakeholders in the defense, commercial, and academic sectors, pointing out the future directions in UGV advancements.

Key Words : Text Mining, Linear Regression Analysis(선형회귀분석), Unmanned Ground Vehicles(UGV), LDA(잠재 디리클레 할당)

#### 기호 설명

$\alpha$  : 문서 내 주제 분포가 얼마나 균일한지 조절

$\beta$  : 주제 내 단어들의 분포가 얼마나 균일한지 조절

$\Theta_d$  : 각 문서  $d$ 가 다루고 있는 주제들의 비율

\* Corresponding author, E-mail: jhjeon@gnu.ac.kr

Copyright © The Korea Institute of Military Science and Technology

$z_{d,n}$  : 문서  $d$ 의  $n$ 번째 단어가 포함됨  
 $w_{d,n}$  : 실제 관측된 문서  $d$ 의  $n$ 번째 단어  
 $\Theta_k$  : 주제  $k$ 로 구성된 단어들의 확률 분포

## 1. 서론

21세기는 군사 기술 혁신이 전쟁 및 전략의 패러다임을 근본적으로 바꾸고 있다. 특히 4차 산업혁명과 연관된 첨단 기술들이 군사 전략에 중요한 영향을 미치고 있는데, 이러한 변화의 핵심에는 무인 시스템의 도입이 있다. 최근 러시아-우크라이나 전쟁에서 무인 항공기(UAV : Unmanned Aerial Vehicle)가 ‘게임 체인저(Game Changer)’로서의 역할을 수행하면서, 기존의 인력 중심 전쟁 양상이 인간과 무인 시스템의 복합 체계로 변모하고 있다는 것을 명확히 보여주고 있다.

이처럼 무인 시스템으로의 전환은 피할 수 없는 추세로 자리잡고 있으며, 무인항공기에 이어 다음으로 주목받는 분야는 바로 무인지상차량(UGV : Unmanned Ground Vehicle)이다. 현재까지 무인항공기는 주로 정찰 역할을 수행하여 유인 전차의 작전을 지원해 왔지만, 무인지상차량의 개발과 전력 통합이 진행되면서, 이들이 유인 차량의 역할을 대체할 가능성이 커지고 있다<sup>[1]</sup>. 무인지상차량은 무인항공기와 더불어, 원격 운용이나 자율 제어를 통해 인간이 탑승하지 않고도 다양한 임무를 수행할 수 있다. 처음에 지뢰 제거를 목적으로 개발된 무인지상차량은 현재 소형 정찰, 무인 전투, 수색, 다목적 병참 지원 등 다양한 목적으로 활용되고 있다. 이 기술은 특히 생화학적 위협, 방사능 오염 지역, 극한의 기후 및 지형 조건 등 인간이 접근하기 어려운 환경에서 중요한 역할을 한다.

세계 무인지상차량 시장은 2021년 미화 31억 달러(약 4조원)에서 2030년 미화 56억 달러(약 7.3조원)으로 확대될 전망이며, 이에 세계 주요 국가들은 무인지상차량의 중요성을 인식하고, 수송, 전투, 정찰 등에 특화된 군용 지상무인차량의 개발과 통합에 박차를 가하고 있다<sup>[2]</sup>.

이렇듯 여러 국가들이 무인지상차량의 전력화 및 유·무인 복합체계로의 전환을 추진 중인 상황에서, 기술의 미래 발전 방향과 전략적 중요성을 이해하는 것은 점점 중요해지고 있다. 이 연구는 국방 분야의 정책 결정자와 연구자들에게 중요한 기초 자료를 제공하며, 미래 전장 환경에 적응하는 효율적인 전략 수립

에 기여할 것이다.

이러한 배경에서, 본 연구는 2004년부터 2020년까지의 글로벌 무인지상차량 관련 특허 데이터를 분석하여, 기술 발전의 주요 추세와 유망 기술, 그리고 잠재적으로 쇠퇴할 기술을 식별한다. 차별적으로 이 연구는 기존의 텍스트 분석을 발전시키는 방법으로 토픽 모델링의 ‘잠재 디리클레 할당(LDA : Latent Dirichlet Allocation)과 선형회귀분석(Linear Regression Analysis)을 결합하여 수행된다. 이 접근법으로 텍스트 마이닝 결과의 신뢰성을 통계적으로 검증하고, LDA로 추출된 각 토픽들의 분포를 선형회귀분석의 변수로 활용하여 각 토픽의 실제 영향력을 정량적으로 평가한다. 이 과정은 텍스트 데이터에서 추출된 정보의 신뢰도를 향상 시키며, 연구 결과의 견고성을 보장할 수 있을 뿐만 아니라 실용적인 의사결정을 위한 유용한 근거를 제공한다. 또한, 이 방법은 무인지상차량의 글로벌 개발 동향을 더욱 명확하게 이해하고, 특히 국방 분야에서 정책 결정자와 연구자들에게 미래 전략을 수립하는 데 필요한 중요한 데이터를 포괄적으로 제공할 수 있다.

## 2. 이론적 배경

무인지상차량은 국방과학기술용어사전(2021)에 정의된 기동무기체계의 한 분야로서, 원격 운용이나 자율 제어에 의하여 사람이 탑승하지 않고 임무를 수행할 수 있는 차량을 의미한다<sup>[3]</sup>. 이는 현대 전장에서 점점 중요해지고 있는 기술로, 군사 및 민간 분야에서 다양하게 활용되고 있다. 급속한 과학기술 발전으로 무인지상차량은 미래 전쟁에서 인명 위험을 줄이면서 효과적인 군사력을 행사하는 데 중요한 역할을 할 것이다. 특히, 지상전에서 무인지상차량의 활용은 인명 피해를 최소화하고 전투 효과를 극대화할 것으로 예상된다. 따라서, 본 연구의 주된 목적은 토픽 모델링(Topic Modeling) 기법을 활용하여 무인지상차량의 기술 트렌드를 분석하고, 향후 미래에 확보해야 할 기술들을 탐색하는 것이다.

### 2.1 무인지상차량 관련 연구

이 분야의 여러 연구들은 국가별 무인지상차량 기술개발 현황, 자율주행 기술, 상호운용성, 센서 기술 및 자율주행 차량 시스템의 미래에 대한 다양한 측면

을 탐구하고 있다. 박승(2009)의 연구에서는 주요 국가들의 무인체계 기술개발 현황을 깊이 있게 분석하였다. 특히 1998년부터 2008년까지의 시기동안 미국, 일본, 유럽, 캐나다, 이스라엘에서 출원된 특허를 주요 분석 대상으로 하였고 이 연구를 통해 각 국가별 기술 수준, 역점을 둔 분야, 그리고 공백 기술 등을 파악할 수 있었다<sup>4)</sup>. 이진호 등(2011)은 무인지상전투차량의 자율주행 기술의 전략적 중요성을 강조하였고, 국내외의 관련 기술 개발 정책을 조사하기 위해 특허와 연구 논문들을 분석하였다<sup>5)</sup>.

### 2.2 토픽모델링과 LDA 알고리즘

텍스트 마이닝은 비정형 텍스트 데이터에서 유용한 정보를 추출하는 과정으로, 토픽모델링은 이 중 하나의 방법이다. LDA(Latent Dirichlet Allocation, 잠재 디리클레 할당)는 토픽 모델링에서 널리 사용되는 방법론으로, 문서 집합 내의 잠재적인 주제 구조를 밝혀내는 데 효과적이다. LDA는 문서가 다양한 주제들의 혼합으로 구성되어 있다고 가정하며, 이를 통해 문서 내 단어들이 특정 주제에 속할 확률을 추정한다<sup>6)</sup>(Blei, Ng, & Jordan, 2003). 본 연구에서 사용된 LDA 알고리즘은 각 문서를 여러 토픽의 혼합으로 표현하고, 각 토픽은 특정 단어들의 분포를 나타낸다. 또한, 깁스 샘플링(Gibb Sampling)을 통한 파라미터 추정을 이용하여, 문서 내에서 각 토픽이 단어를 생성하는 비율을 계산하고, 이를 바탕으로 문서의 주제를 분류한다<sup>7)</sup>(G. Casella, & E. George, 1992). 이 방법은 주제별 핵심 단어를 식별하고 텍스트 데이터로부터 의미 있는 패턴을 추출하는데 유용하다. Fig. 1은 이러한 LDA 알고리즘을 설명한 것이다.

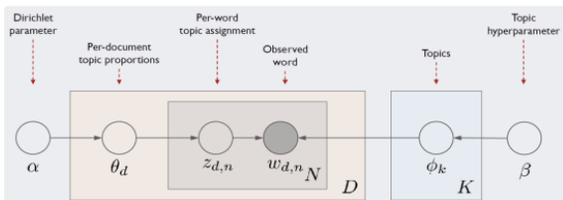


Fig. 1. Graphical model of LDA(Blei, Ng, & Jordan, 2003)

### 2.3 토픽 모델링 관련 연구

기술 동향을 파악하기 위해 논문, 특허, 그리고 동향 분석 보고서 등의 자료가 널리 사용되지만, 이러한

방대한 양의 데이터에서 정량적이고 구조화된 정보를 추출하는 것은 어려운 과제이다. 이에 대한 해결책으로, 텍스트 마이닝의 한 분야인 토픽 모델링 기법을 활용한 연구가 활발히 진행되고 있다. 박건철 등(2019)은 스마트시티 연구의 주요 토픽과 동향을 분석하기 위해 11,527건의 학술 논문을 토픽 모델링으로 조사했다<sup>8)</sup>. 전고운 등(2017)은 국방 분야에서 장갑전투차량을 대상으로 한 토픽 모델링을 사용하여 무기체계 연구개발 동향을 분석했다<sup>9)</sup>. 류병준 등(2023)은 한국, 미국, 일본, 중국을 중심으로 한 해양무인체계의 연구 동향을 토픽 모델링을 이용한 특허 분석으로 조사했다<sup>10)</sup>. 백슬아 등(2023)은 UAM(도심항공교통)분야의 연구주제 변화 추세를 다이나믹 ‘토픽 모델링(DTM, Dynamic Topic Modeling)’으로 분석했으며<sup>11)</sup>, 김은정 등(2023)은 헬스케어 분야의 뉴스 빅데이터를 토픽 모델링으로 분석하여 산업의 주요 이슈와 동향을 파악했다<sup>12)</sup>. 김미라(2013)의 연구에서는 한국 해양, 해군, 수중 군사과학 분야의 트렌드를 파악하기 위해, 1998년부터 2012년까지 「한국군사과학기술학회지」에 게재된 논문들의 영어 키워드를 분석했으며, 총 454개의 키워드와 107편의 논문을 분석 대상으로 하여 해당 분야의 미래 발전 방향과 추가 연구의 필요성을 제시하였다<sup>13)</sup>.

## 3. 연구 방법

### 3.1 연구설계

본 연구는 무인지상차량 기술 동향을 분석하는 데 중점을 두고 설계되었다. 연구의 핵심은 국가별 무인지상차량 관련 특허에 대한 분석으로, 핵심기술, 세부 기술에 대한 특허 동향을 탐색하여 기술과 연구 트렌드를 파악하는 것이다. 이 연구는 Python 3.8.1을 사용하며 Google Colab. 환경에서 실행되었다. 연구 절차는 아래 Fig. 2와 같다.

#### 3.1.1 데이터 수집

데이터 수집 단계에서는 KIPRIS(Korea Intellectual Property Rights Information Service, 한국지식재산권정보서비스)를 활용하여 무인지상차량 관련 핵심 단어들을 참조하여 국가별 특허를 수집하였다. 이 과정에서 IP5(Intellectual Property)에 해당하는 주요 5개국 즉 미국, 유럽, 중국, 일본, 한국의 특허 데이터를 집중적

으로 조사하였다. IP5는 세계의 주요 지식재산권 시장을 대표하는 국가들로, 이들 국가의 특허 데이터는 해당 분야의 글로벌 혁신 동향을 이해하는 데 중요한 역할을 한다. 수집된 데이터에는 각 단계에서 공개된 특허뿐만 아니라, 심사를 통과하여 등록된 특허도 포함되어 있다. 국가별로 동일한 검색 조건을 적용하여 2004년부터 2020년까지의 기간동안 총 10,309건의 특허 데이터를 수집했다. 이러한 데이터 수집 과정은 본 연구의 주제인 무인지상차량 기술 발전 동향을 분석하기 위한 기초 자료로 활용되었다. Table 1을 통해 국가별로 수집된 특허 출원 수를 나타내었다.

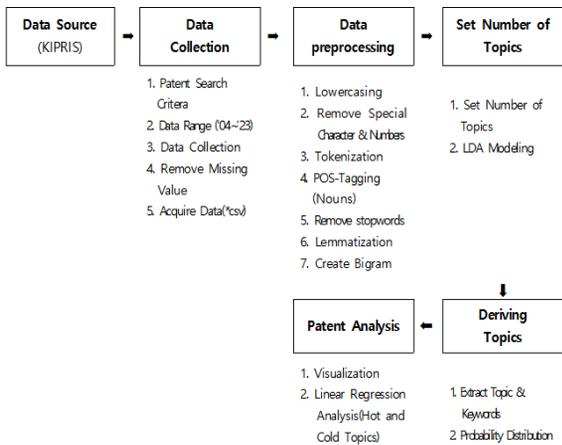


Fig. 2. Flowchart

Table 1. Number of patent application

| Country | Patent | Portion |
|---------|--------|---------|
| US      | 4,052  | 39.3 %  |
| EU      | 3,913  | 38 %    |
| China   | 1,232  | 12 %    |
| Japan   | 955    | 9.2 %   |
| Korea   | 157    | 1.5 %   |
| Total   | 10,309 | 100 %   |

### 3.1.2 데이터 전처리

데이터 전처리 단계에서는 특허 데이터의 제목과 요약 부분에 대한 전처리 작업을 수행하였다. 이 과정에서 텍스트 데이터의 특성을 고려하여 품사 결정(POS-

Tagging)을 통한 명사 추출, 원형 복원(Lemmatization)과 2개의 연속된 단어(Bi-gram)을 활용하였다.

### 3.1.3 LDA 모델 분석

토픽 모델링 단계에서는 전처리된 데이터를 바탕으로 LDA(Latent Dirichlet Allocation) 모델을 통해 분석을 진행하였다. 이 단계에서는 토픽의 개수를 선정하는 것이 중요한 과정이며, 적절한 토픽 수의 선정은 분석의 정확성과 유용성을 결정짓는 중요한 과정이다<sup>[4]</sup>(T. Griffiths, & M. Steyvers, 2004). 이를 위해 ‘응집도 지수(Coherence Score)’를 활용하였다.

### 3.1.4 LDA 모델 선정 이유

토픽 모델링 방법론은 다양하며, LDA 외에도 ‘비음수 행렬 분해(NMF : Non-negative Matrix Factorization)’과 ‘확률 잠재 의미 분석(pLSA : Probabilistic Latent Semantic Analysis)’ 등 여러 기법이 존재한다. 본 연구에서 LDA를 선택한 주된 이유는, Table 2에서 나타난 바와 같이 여러 우수한 특징을 갖추고 있기 때문이다.

우선, LDA는 명확한 주제 해석을 가능하게 하며 확률 모델을 기반으로 하여 선형회귀분석과의 통합이 용이한 장점이 있다. 반면에 NMF와 pLSA는 확률 모델을 기반으로 하지 않거나 선형회귀와의 통합이 어렵다. 또한, LDA는 토픽의 확률을 계산하여 각 문서에 대한 토픽 분포를 제공하며 이는 선형회귀 모델과의 통합시 중요한 요소로 작용한다. 이러한 측면을 고려할 때, LDA는 본 연구 목적과 필요조건에 가장 가장 부합하는 최적의 토픽 모델링 방법론으로 판단된다.

Table 2. Comparison of topic modeling methods

| 기준        | LDA       | NMF    | pLSA      |
|-----------|-----------|--------|-----------|
| 명확한 주제 해석 | 명확        | 어려움    | 어려움       |
| 선형회귀분석 통합 | 가능        | 어려움    | 어려움       |
| 확률 모델링 기반 | O         | X      | O         |
| 활용 분야     | 토픽 모델링 전반 | 이미지 분석 | 문서, 정보 검색 |
| 계산 효율성    | 보통        | 빠름     | 느림        |

### 3.1.5 주제 도출

이 연구의 마지막 주제 도출 단계인 주제 도출에서는, 시각화와 통계적 방법을 적용하여 토픽 모델링 결과를 분석하였다. 이를 통해 핵심 주제 선정, 국가별 핵심 기술 비교, ‘Hot’, ‘Neutral’ 그리고 ‘Cold’ 토픽을 선정하고 무인지상차량 기술 분야의 방향성에 대한 통찰을 얻고자 하였다.

## 4. 연구 결과

### 4.1 특허 데이터 분석

4장 연구 결과에서는 3장의 연구 방법을 통해 2004년부터 2023년까지 IP5 5개국(미국, 유럽, 중국, 일본, 한국)의 무인지상차량 관련 특허 출원 동향을 조사하고, 이를 통해 각 국가의 기술 발전 양상을 분석하였다. 특히, Table 1에 나타난 주요 국가별 특허 출원 비율에 따르면, 미국(39.3%)과 유럽(38%)이 전체 출원 건수의 대부분을 차지하며 이 분야에서 선두적인 역할을 하고 있는 것으로 나타났다. 이는 이들 지역이 무인지상차량 기술 발전에 있어서 중추적인 역할을 하고 있음을 시사한다.

그러나 이 연구에서는 2021년부터 2023년까지의 데이터를 배제하고 분석을 진행하였다. 그 이유는, Fig. 3에서 확인할 수 있듯이 이 기간동안 IP5 국가들의 특허 출원 건수가 급격하게 감소하는 경향이 관찰되었기 때문이다. 이러한 감소 추세는 잠재적으로 COVID-19 팬데믹으로 인한 경제적, 운영적 어려움, 기술 발전의 일시적 정체 또는 특허 정책 변화 등 다양한 외부 요인들이 영향을 미쳤을 가능성이 있으며, 보통 미국, 유럽, 한국 특허청의 경우 특허 출원 후 약 18개월 이내에 공개되는 특허들이 포함되어 있어 최근 출원된 특허의 정보가 아직 공개되지 않았을 수도 있다는 점도 고려하였다. 이 시기의 데이터는 무인지상차량 기술 발전의 장기적 추세와 패턴에 대한 오해나 왜곡된 정보를 제공할 수 있다고 판단되었다. 이러한 비정상적인 시기의 데이터를 배제하고, 2004년을 시작으로 2020년까지의 더 안정적인 기간의 데이터를 기반으로 연구를 수행하였다.

아래 Fig. 4의 그래프는 2021년부터 2023년까지의 데이터를 제외한 분석 결과를 보여준다. 국가의 연도별 특허 출원 건수에서 보이는 추세에 따르면, 미국은 특허 2015년 이후 특허 출원 수가 급증하며 기술 발

전에 있어서 증가세를 보였다. 유럽 역시 2016년부터 출원 수가 급격히 증가하여, 무인지상차량 기술의 빠른 성장과 발전이 진행되고 있음을 나타낸다. 반면, 중국, 일본 및 한국은 상대적으로 낮은 비율을 유지하고 있으나, 중국은 특히 2012년 이후 눈에 띄는 출원 증가세를 보여 시장에서의 입지를 확대하고 있음을 시사한다.

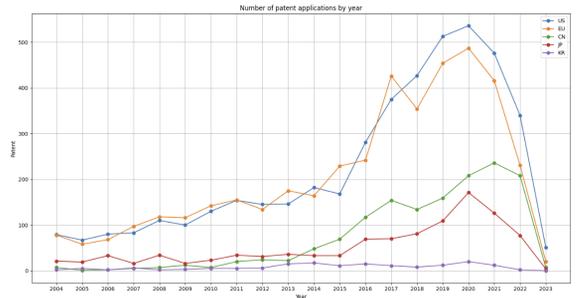


Fig. 3. Annual patent application trend('04~'23)

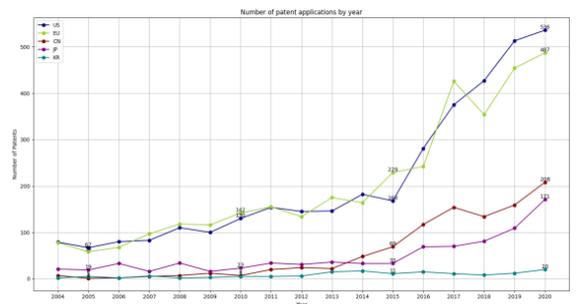


Fig. 4. Annual patent application trend('04~'20)

### 4.2 Bi-gram을 활용한 데이터 전처리

데이터 전처리 단계에서 본 연구는 특허 데이터의 제목과 요약 부분에 대한 전처리 작업을 실시하였다. 이 과정에서 텍스트 데이터의 특성을 고려하여 형태소 분석을 통한 명사 추출 및 원형복원과 두 개의 연속된 단어, 즉 ‘Bi-gram’을 활용하였다. 단일단어인 ‘Uni-gram’로 전처리한 결과, ‘System’, ‘Control’, ‘Data’, ‘ground’, ‘power’와 같은 단어들도 도출되었으나 이들 단어만으로는 기술에 대한 심층적인 분석을 진행하기에는 충분하지 않았다. 따라서, ‘Bi-gram’을 활용하여 ‘Control\_system’, ‘Data\_transmission’, ‘Image\_sensor’, ‘power\_supply’와 같은 기술적으로 의미 있는 복합 단어를 도출할 수 있었다. 결과적으로, ‘Bi-gram’을 사

용한 토픽 모델링은 기술적 맥락이 더욱 뚜렷한 키워드를 추출하여 무인지상차량 분야에서의 기술 개발 동향을 파악하는 데 효과적이었다.

#### 4.3 최적 토픽 수 선정

토픽 모델링 분석에서 가장 중요한 요소는, 분석에 적합한 토픽의 수를 선정하는 것이다. 본 연구에서는 최적의 토픽 수를 결정하기 위해 ‘응집도 지수(Coherence Score)’를 활용하였다. 이는 토픽 모델링 결과의 질을 평가하는 지표로서, 각 토픽 내 단어 사이의 의미론적 일관성을 측정한다. 높은 응집도 지수는 토픽이 잘 구성되어 있음을 의미하며, 여러 연구에서 토픽 모델링 결과를 평가하는 데 성공적으로 사용되었다<sup>[8-10]</sup>. 아래 Fig. 5에 제시된 응집도 지수 그래프를 분석한 결과, 8개의 토픽일 때 변동성이 크고 응집도가 낮음을 확인할 수 있었고, 12개의 토픽일 때는 변동성이 가장 낮고 응집도가 가장 높은 값을 나타내었다. 이는 12개의 토픽이 데이터 내의 다양한 주제들을 가장 잘 반영하고 있으며, 각 토픽 간의 구분이 잘 이루어지고 있음을 의미한다. 그러므로, 최적 토픽 수를 12개로 결정하였다.

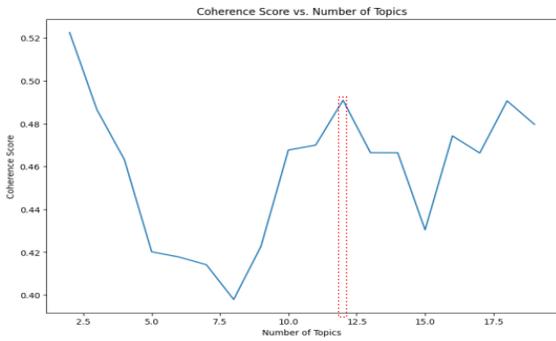


Fig. 5. Coherence score

#### 4.4 토픽 모델링 분석 및 핵심 기술 선정

이 연구에서는 수집된 10,309건의 무인지상차량 관련 특히 데이터를 대상으로 토픽 모델링을 수행하였으며, 이를 통해 해당 기술 분야의 연구개발 동향과 핵심 주제들을 탐색하는 데 중점을 두었다. Fig. 6의 ‘워드 클라우드(Word Cloud)’는 각 토픽에 속한 주요 키워드들의 등장 빈도와 중요성을 시각적으로 잘 나타내고 있으며, 이는 복잡한 데이터셋의 구조를 단순화하여 연구 결과를 직관적으로 전달한다.

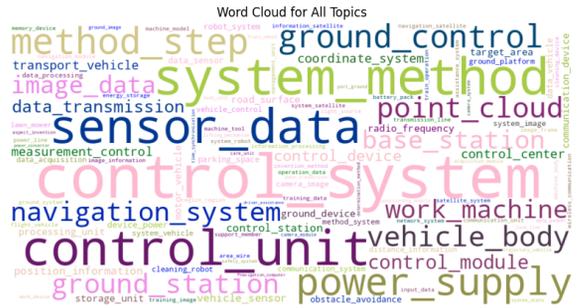


Fig. 6. Word cloud for all topics

또한, Table 3은 앞서 수행한 종합적인 토픽 모델링 과정에서 도출된 키워드들의 상호 연관성과 각 토픽별 주제의 정보적 가치를 분석하여, 무인지상차량 기술의 주요 트렌드를 종합적으로 보여주는 통계적 요약 을 제공한다. Topic 2의 결과는 아래 Fig. 7에서 키워드별 비중의 순서대로 나타나 있으며, ‘control\_system (제어 시스템)’, ‘control\_unit(제어 유닛)’, ‘power\_supply (전력 공급)’ 등의 키워드가 높은 비중을 차지함으로써, 해당 토픽의 중요성을 강조한다. 이러한 핵심 키워드들은 토픽의 명확한 정의를 형성하는 데 기여하며, 이를 바탕으로 ‘Control system and power supply (제어 시스템 및 전력 공급)’이라는 주제 명칭을 선정 하였다. 이 주제명은 방위산업체에서 30년 이상 기계·제어 분야 R&D 경력을 보유한 전문가의 깊은 자문을 통해 이루어졌다. 이 전문가는 전자 제어 시스템에 관한 지식을 담은 ‘카 일렉트로닉스 교본’의 저자로, 그의 풍부한 경험과 전문 지식은 주제 선정 과정에서 핵심적인 역할을 했다. 특히 복잡한 기계 및 전자 제어 시스템에 대한 깊은 이해를 바탕으로, 토픽 모델링을 통해 도출된 특정 기술 영역들과 무인지상차량 기술 발전에 있어 중요한 요소들과의 연관성을 강조하며 본 주제 선정 과정에 기여했다.

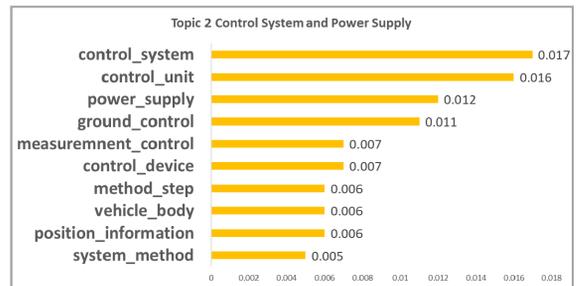


Fig. 7. Portion distribution of keywords in topic 2

Table 3. Topic modeling analysis result of patents

| No. | Subject  | Keyword   | Portion       |
|-----|--|---|---------------|
| T1  | Vehicle and Road Environment Interaction<br>(차량과 도로 환경 상호작용) | vehicle_body, road_surface, point_cloud, parking_space, communication_system, communication_unit, communication_device, wireless_communication, battery_pack, care_unit           | 0.0842        |
| T2  | Control System and Power Supply<br>(차량 제어 및 전력 공급)           | <b>control_system, control_unit, power_supply, ground_control, control_device, measurement_control, position_information, vehicle_body, method_step, system_method</b>            | <b>0.1837</b> |
| T3  | Control for Military Satellite<br>(군사 위성 지상 제어)              | control_station, satellite_system, data_base, data_acquisition, navigation_satellite, system_satellite, information_satellite, ground_unit, transmission_line, input_data         | 0.0645        |
| T4  | Ground Data Communication<br>(지상 데이터 통신)                     | ground_station, base_station, data_transmission, control_module, system_method, coordinate_system, method_step, control_system, processing_unit, navigation_system                | 0.0986        |
| T5  | Mechanized Equipment<br>(기계화 장비)                             | machine_tool, machine_model, region_model, memory_device, front_wheel, lane_line, acquisition_module, state_system, ground_image  | 0.0472        |
| T6  | RF Communication and Target Recognition<br>(무선 통신 및 표적 인식)   | radio_frequency, target_area, distance_information, information_processing, power_line, management_unit, camera_system, host_cell, aspect_invention, camera_module                | 0.0505        |
| T7  | Transport and Vehicle Control System<br>(운송용 차량 제어 시스템)      | control_system, transport_vehicle, methods_step, method_system, system_image, vehicle_control, camera_image, image_information, system_robot, light_source                        | 0.0771        |
| T8  | Ground Working and Safety<br>(지상 작업 및 안전 시스템)                | ground_system, cleaning_device, operation_data, lifting_mechanism, determination_method, data_packet, vehicle_trajectory, time_synchronization, positioning_module, safety_system | 0.0504        |
| T9  | Ground Control System<br>(지상 제어 시스템)                         | ground_control, work_machine, control_center, storage_unit, ground_device, train_operation, control_unit, energy_storage, work_device, part_ground                                | 0.0582        |
| T10 | Navigation and Energy Supply System<br>(항법 및 전력공급 시스템)       | navigation_system, power_supply, obstacle_avoidance, device_power, flight_vehicle, network_system, area_wire, navigation_module, navigation_computer, support_member              | 0.0726        |
| T11 | Sensor Data and Image Analysis<br>(센서 데이터 및 영상 분석)           | sensor_data, image_data, system_method, vehicle_sensor, data_sensor, data_vehicle, data_acquisition, system_vehicle, ground_platform, data_processing                             | 0.1193        |
| T12 | Equipment Assistance System<br>(장비 보조 시스템)                   | assistance_system, machine_body, driver_assistance, training_image, training_data, system_method, image_frame, power_transmission, lawn_mower, power_converter                    | 0.0672        |

#### 4.5 Hot and Cold Trend 분석

본 연구에서 수행된 토픽 모델링은 시간이 흐름에 따라 각 토픽의 중요성이 어떻게 변화하는지에 대해

동적인 흐름을 보여준다. 2004년부터 2020년까지 각 토픽의 연간 비중 변화를 추적함으로써, 연구 및 개발의 중요성이 증가하거나 감소할 것으로 예상되는 기

술을 식별하는 과정을 수행하였다. 아래 Fig. 8은 이러한 결과를 그래프로 나타내었다. 이 연구에서 수행한 선형회귀분석은 각 토픽의 연도별 비중을 ‘종속변수(y)’, 시간을 ‘독립변수(x)’로 설정하여 토픽의 중요성 변화를 평가하는 데 사용되었다.

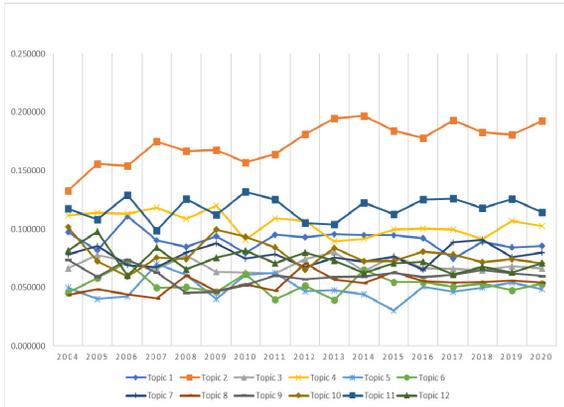


Fig. 8. Annual trend of topic proportion

분석의 첫 단계인 ‘회귀계수(Coefficient)’ 분석에서 Fig. 9는 ‘Topic 2’가 양의 회귀계수를 가지고 있음을 보여주며, 시간이 지날수록 해당 토픽의 비중이 증가하고 있음을 나타낸다.

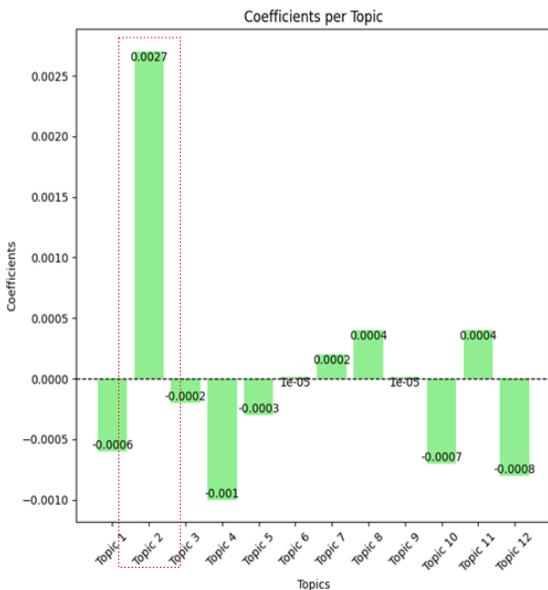


Fig. 9. Regression coefficient per topics

다음으로, Fig. 10에서 Topic 2의 ‘유의 확률(P-value)’이 0.05 이하로 나타나, 이 증가 추세가 통계적으로 유의미하다는 것을 확인시켜 준다. 또한 Fig. 11에서 Topic 2의 ‘결정계수(R-squared values)’도 다른 토픽들에 비해 월등하게 높은 수치인 0.663으로 나타났다. 이러한 결과는 ‘Topic 2’를 ‘Hot Topic’으로 분류하는 강력한 근거를 제공한다, 특히 Fig. 12에서는, ‘Topic 2(제어 시스템 및 전력 공급)’의 비중이 분석 대상 기간인 2004년부터 2020년까지 지속적으로 증가하는 경향을 보여준다. 이러한 경향은 ‘Topic 2’가 미래 기술 발전에서 중요한 역할을 할 것으로 예상되는 유망기술로 선정될 수 있음을 시사한다.

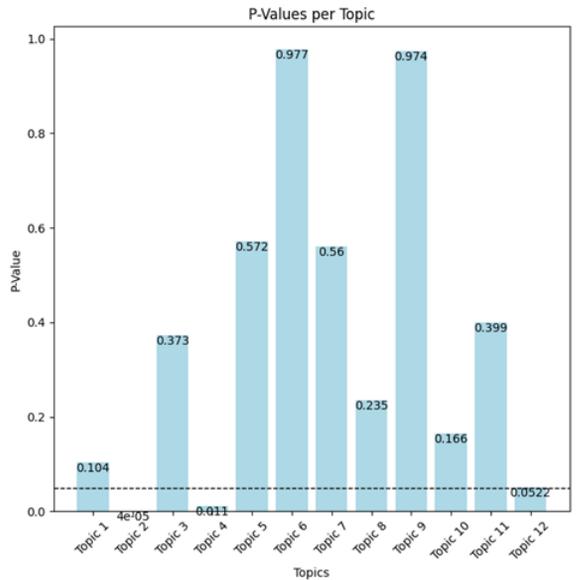


Fig. 10. Regression p-value per topics

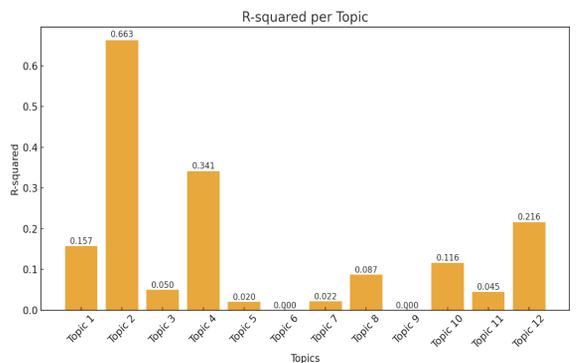


Fig. 11. Regression R-squared per topics

반면에 회귀계수가 음수이고, P-value가 0.05 이상인 토픽들은 ‘Cold Topic’으로 분류되었다, 대표적인 예로, Fig. 13에 나타난 ‘Topic 1(Vehicle and Road Environment Interaction)’은 시간에 따라 그 비중이 감소하는 추세를 보여준다. 이러한 추세는 해당 주제의 연구 비중 감소나 중요성 저하가 예상되는 분야

로 해석된다. 이와 유사하게, ‘Topic 5(Cleaning Robot and Equipment)’, ‘Topic 8(Operation Data and Ground System)’ 또한 이 범주에 포함된다.

이외에도, ‘Neutral Topic’으로 분류된 주제들은 회귀계수가 음수이지만 P-value가 유의수준에 근접한 것으로 나타난다. 이는 해당 주제들이 아직 명확한 발전

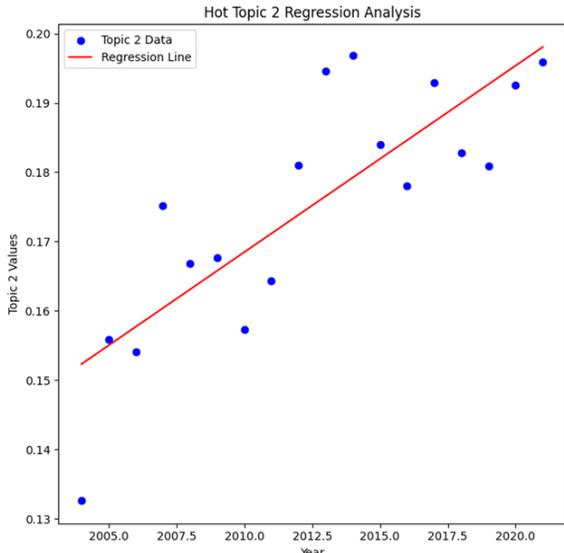


Fig. 12. Regression analysis of hot topics

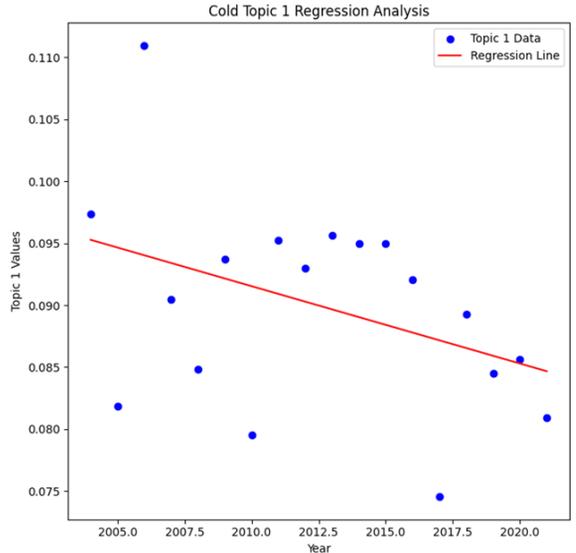


Fig. 13. Regression analysis of cold topics

Table 4. Hot and cold topic

| Hot & Cold | Topic | Subject                                  | Coefficient | P-value      |
|------------|-------|--|-------------|--------------|
| Hot        | T2    | Control System and Power Supply          | 0.0027      | 0.0000391*** |
| Neutral    | T4    | Ground Data Communication                | -0.0001     | 0.011*       |
|            | T12   | Equipment Assistance System              | -0.0008     | 0.0522       |
| Cold       | T1    | Vehicle and Road Environment Interaction | -0.0006     | 0.104        |
|            | T3    | Ground Control for Satellite             | -0.0002     | 0.373        |
|            | T5    | Cleaning Robot and Mechanized Equipment  | -0.0003     | 0.572        |
|            | T6    | RF Communication and Target Recognition  | -0.00001156 | 0.977        |
|            | T7    | Transport and Vehicle Control System     | 0.0002      | 0.56         |
|            | T8    | Ground Working and Safety                | 0.0004      | 0.235        |
|            | T9    | Ground Control System                    | 1.137E-05   | 0.974        |
|            | T10   | Navigation System and Energy Supply      | -0.0007     | 0.166        |
|            | T11   | Sensor Data and Image Analysis           | 0.0004      | 0.399        |

\* P<.05, \*\* P<.01, \*\*\* P<.001

혹은 쇠퇴 추세를 보이지 않고, 향후 연구개발의 방향에 따라 중요성이 변할 수 있음을 시사한다. 또한, 통계적 유의성이 결정적이지 않음에도 불구하고, 이러한 주제들이 미래의 연구와 개발에 중요한 영향력을 발휘할 가능성을 내포하고 있음을 의미한다. 이는 ‘Neutral Topic’ 분류가 미래의 연구 동향과 전략적 결정에 중요한 정보를 제공할 수 있음을 시사한다. 이러한 종합적인 결과를 정리한 표는 Table 4와 같다.

## 5. 결론

이 연구는 무인지상차량 관련 특허 데이터를 분석하여 글로벌 연구 동향을 탐색하고, 이를 바탕으로 미래의 유망 기술과 잠재적으로 쇠퇴할 기술을 식별하는 데 중점을 두었다. 토픽 모델링과 선형회귀분석을 활용하여 각 주제의 시간에 따른 연구 비중의 변화를 추적하고, 이를 통해 연구개발의 중요성이 증가하는 추세에 있는 기술을 ‘Hot Topic’으로, 중요성이 감소할 가능성이 있는 기술을 ‘Cold Topic’으로 분류하였다. 그 결과 ‘제어 시스템 및 전력공급’ 주제가 높은 유의성을 가지며 시간이 지남에 따라 중요성이 증가하는 경향을 보인다는 점이다. 이는 해당 기술이 미래에 중요한 연구 분야로 자리매김할 가능성이 높음을 시사한다.

본 연구는 글로벌 트렌드를 포괄적으로 분석함으로써, 무인지상차량 기술의 발전을 넓은 관점에서 이해할 수 있는 기반을 제공한다. 시간적 추세 분석을 통한 유망 기술의 예측은 장기적인 기술 발전 추세를 이해하고, 적절한 연구개발 전략을 수립하는 데 중요한 도구가 될 수 있다. 또한, 이 연구는 정책 입안자와 연구자에게 연구 및 개발 투자에 있어 중요한 가이드라인을 제공하며, 연구개발 전략을 맞춤형으로 수립하는 근거를 제공할 수 있기를 기대한다. 그러나 이와 같은 시사점에도 불구하고 다음과 같은 한계점을 가지고 있다.

첫 번째, 연구에 활용된 무인지상차량 관련 특허 데이터가 기술 보안으로 인해 외부에 공개되지 않아 제한된 데이터로 방산 분야에서의 무인지상차량의 중요한 기술 동향을 파악하는 데는 한계가 있었다. 향후 연구에서는 보다 다양한 데이터 소스와 접근 방법을 활용하여 데이터 수집 범위를 확대할 수 있다면 방산 분야에서의 기술 동향 분석의 품질이 향상될 것이라

생각한다.

두 번째, 토픽 모델링을 통해 도출된 결과는 보다 구체적으로 세부적인 분류와 범위 축소가 제한되었다. 특히 ‘Hot Topic’으로 분류된 ‘제어 시스템 및 전력공급’과 같은 주제에 대해서도, 기술의 구체적인 하위 단계에 대한 세분화가 어려웠다. 이는 토픽 모델링 기법의 한계로서 주제의 세부적인 특성과 변화를 파악하는 데 제약이 있다는 것을 알게 되었다. 향후 연구에서는, 더욱 정교하고 다양한 분석 기법과 데이터 소스를 활용하여 이러한 한계를 보완하고, 보다 세부적이고 구체적인 기술 동향 분석 연구를 수행할 수 있을 것이다.

세 번째, 본 연구에서 일반적인 무인지상차량(UGV) 특허 동향을 분석한 결과와 함께, 국가별 무인지상차량 정책 및 기술상황을 보다 심층적으로 연구하고 이를 트렌드 분석 결과와 비교했다면 더욱 풍부한 인사이트를 제공할 수 있었을 것이다. 향후 연구에서는 이러한 접근을 포함하여 각 국가별로 중점을 두고 있는 기술들을 분석하고, 이러한 기술들이 글로벌 트렌드와 어떻게 상호작용하는지를 조명할 예정이다. 이는 연구 결과의 실용성을 바탕으로, 국가별 연구개발 전략을 맞춤형으로 수립하는 데 필요한 근거를 제공할 수 있을 것으로 기대한다.

## 후 기

이 성과는 2019년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2019R1A2C1090655). 또한 산업통상자원부의 ‘융합기술사업화 확산형 전문인력 양성 사업’의 지원을 받아 수행된 논문임.

그리고 본 연구는 제1저자의 석사학위논문을 기반으로 작성한 것임.

## References

- [1] K. Kim, T. Kim, “Operational Concepts and Development Directions for Unmanned Combat Vehicles in Future Battlefields,” Journal of the Korea Defense Industry Association, Vol. 527, pp. 130-139, 2023.

- [2] S. Kim, "Development Trends of Military Unmanned Ground Vehicles," *Industrial Trends*, Vol. 97, pp. 1-3, July, 2022.
- [3] Dictionary of National Defense Science and Technology Terms, "Unmanned Ground Vehicles (UGV)," Korea Research Institute for defense Technology planning and advancement, 2021.
- [4] S. Park, "A Study on Trend of Technology Development for Unmanned Combat Ground Vehicle," *Korea Academia-Industrial cooperation Society*, Vol. 10, No. 7, pp. 1735-1739, 2009.
- [5] J. Lee, S. Kim, C. Lee, "The Development Trend Analysis of Autonomous Driving Technology for Unmanned Ground Combat Vehicles," *Journal of the Korea Institute of Military Science and Technology*, Vol. 14, No. 5, pp. 760-767, 2011.
- [6] D. M. Blei, A. Y. Ng, M. I. Jordan, and M. I., "Latent dirichlet allocation," *Journal of machine Learning research*, Vol. 3, pp. 993-1022, 2003.
- [7] G. Casella, Edward I. George, "Explaining the Gibbs Sampler," *The American Statistician*, Vol. 46, No. 3, pp. 167-174, 1992.
- [8] K. Park, C. Lee, "A Study on the Research Trends for Smart City using Topic Modeling," *Journal of Internet Computing and Services(JICS)*, Vol. 20, No. 3, pp. 119-128, 2019.
- [9] G. Jeon, I. Kang, J. Jeon, "Systematic Analysis on the Trend of Defense Technologies Using Topic Modeling : A Case of an Armoured Fighting Vehicle," *The Journal of Industrial Innovation*, Vol. 36, No. 1, pp. 69-94, 2020.
- [10] B. Yu, Y. Ha, "Analysis of Research Trends of Unmanned Marine Systems in Korea, the United States, Japan, and China using Topic Modeling," *Journal of the Korea Academia-Industrial cooperation Society*, Vol. 23, No. 11, pp. 395-403, 2022.
- [11] S. Baek, C. Song, J. Jeon, "Analysis of research trends related to urban air mobility(UAM) through dynamic topic modeling," *Innovation studies*, Vol. 18, No. 1, pp. 119-142, 2023.
- [12] E. Kim, H. Choi, "Analyzing Core Technology and Technological Convergence in Healthcare Using Topic Modeling and Network Analysis: Focus on Patent Information," *Journal of the Korea Institute of Information and Communication Engineering*, Vol. 26, No. 5, pp. 763-778, 2022.
- [13] M. Kim, "Trends Analysis on Marine/Naval/Underwater Military Science and Technology by Keywords Analysis," *Journal of the Korea Institute of Military Science and Technology*, Vol. 16, No. 5, pp. 622-630, 2013.
- [14] T. Griffiths, M. Steyvers, "Finding scientific topics," *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES*, Vol. 101, pp. 5228-5235, 2004.