

SLM 기반의 자연어 음성 인식 처리기술을 활용한 재난구조용 자율 드론 연구

채웅*, 김하늘**, 이병찬**, 이강봉**, 우한별**, 이형주**, 조영호(교신저자)***
*국방대학교 국방관리대학원 국방과학학부 사이버컴퓨터공학과 박사과정
**국방대학교 국방관리대학원 국방과학학부 사이버컴퓨터공학과 석사과정
***국방대학교 국방관리대학원 국방과학학부 사이버컴퓨터공학과 교수
e-mail:younghocho@korea.kr

A Research on Autonomous Drones for Disaster Rescue Using SLM-Based Natural Language Speech Recognition Technology

Woong Chae*, Haneul Kim**, Byungchan Lee**, Kangbong Lee**, Hanbyeol Woo**,
Hyungjoo Lee**, Youngho Cho(Corresponding Author)***
*Ph.D. Course, Dept. of Cyber Security and Computer Engineering,
Korea National Defense University
**Master's Course, Dept. of Cyber Security and Computer Engineering,
Korea National Defense University
***Professor, Dept. of Cyber Security and Computer Engineering,
Korea National Defense University

요약

본 연구에서는 재난상황에서 SLM(Small Language Model) 기반의 자연어 음성 인식 처리기술을 활용하여 구조대상자의 음성 인식과 상황 파악을 통해 효과적으로 재난구조 임무 수행이 가능한 있는 재난구조용 자율 드론을 제안한다. 다양한 분야에서 활용되는 드론에 SLM 기술을 적용하여, 구조대상자와 드론이 상호 대화를 수행하고 이를 통해 드론이 구조대상의 상황과 요구 사항을 정확히 판단하고 필요한 구조임무를 수행하도록 함으로써 복잡하고 위험한 재난상황 속에서 효과적인 구조활동을 수행할 수 있다. 구체적으로, 제안하는 자율 드론의 두 가지 구조상황별 드론 활용 시나리오를 통해 어떻게 제안 드론이 효과적으로 구조 임무를 수행할 수 있는지를 설명한 후, 그러한 임무 수행을 위한 자율 드론 시스템 구성도와 함께 핵심 모듈(음성인식 및 자연어 처리 모듈, 상황평가 및 우선 임무순위 분류 모듈)의 동작 기능을 단계적으로 기술하였다. 이를 통해, 본 연구에서 제안한 자율 드론은 기존의 드론 활용 대비 더욱 효과적인 구조 임무를 수행할 수 있을 것이라 기대한다.

1. 서론

전 세계적으로 드론산업의 성장세는 날이 갈수록 증가하고 있으며, 드론을 활용한 분야 역시 공공 안전, 재난구조, 농업, 시설 점검 및 배송 등 다양해지고 있다[1, 2]. 특히, 건물 붕괴, 산불 화재, 지진 등과 같은 재난구조 상황에서는 사람이 직접 투입하기 어려운 공간적 측면 및 위험 등이 있으므로 드론과 같은 무인 자율체계의 활용이 적극적으로 요구된다.

최근 인공지능(AI) 기술의 발전과 함께 ChatGPT와 같은 LLM (Large Language Model)은 사람보다 더 사람처럼 튜링 테스트를 통과함에 따라 사람과의 의사소통이 충분히 가능해졌다[3, 4]. 또한, SLM(Small Language Model)은 LLM에 비해 크기는 작지만, 그에 필적하는 충분한 성능을 보장하고, 특정 요구사항에 맞춰 운영하는데 효율적이므로 SLM을 드론에 탑재하여 충분히 활용한다면 그에 따른 상당한 효과가 있을 것이다[5, 6, 7].

드론의 음성인식과 관련한 최근 연구 및 산업 동향을 확인한 결과, 드론을 통해 재난현장 내 구조대상자의 말소리, 고향, 박수 등을 인식하여 활용하는 방식의 연구[8]와 단순 음성지시를 통한 움직임에 관한 연구[9, 10] 등을 확인할 수 있었다. 하지만, 복잡하고 혼란스러운 상황 속에서 현장지휘자에 대한 의존도를 낮추고 무인 자율 드론이 주도적인 구조활동을 수행할 수 있는 방안에 대한 추가적인 연구가 필요하다고 판단하였다.

이에 따라, 본 연구에서는 'SLM 기반의 자연어 음성인식 처리 기술을 활용한 재난구조용 자율 드론 기술'을 제안하고 연구하고자 한다. 재난구조 상황에서 구조대상자와 드론이 자연어 음성을 통해 상호 소통하는 가운데, 드론이 SLM 기술을 활용하여 구조대상자의 상태와 요청사항을 정확히 이해한 후 자체적인 구조 계획을 수립한다면 능동적이고, 효과적인 구조 임무 수행이 가능할 것이다.

이후 본 논문의 구성은 다음과 같다. 2장에서는 제안하는 자율

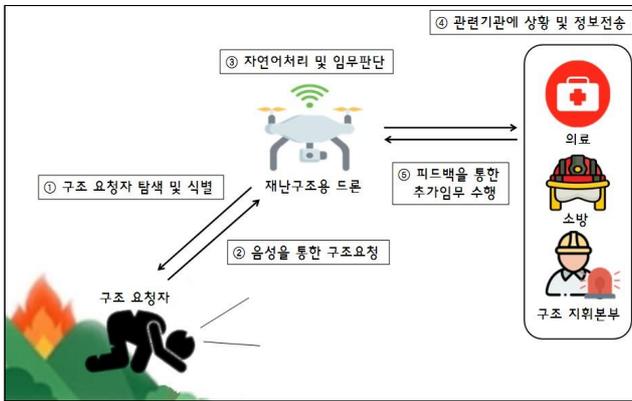
드론의 상황별 활용 시나리오와 드론 시스템 구성을 기술하고, 3 장에서는 결론 및 향후 연구 방향을 제시한다.

2. 상황별 시나리오 및 시스템 구성

2.1 상황별 시나리오

재난구조라는 복잡한 상황 중에서도 자율 드론이 지하 등에서 임무를 수행함에 따라 통신이 단절된 상황과 대량 사상자가 발생하여 다수의 드론이 임무를 수행하는 상황 등 총 두 가지의 시나리오를 통해 본 연구에서 제안하는 자율 드론의 행동 절차에 대해 구체적으로 기술한다.

공통적인 재난구조 상황과 구조 시나리오는 아래의 [그림 1]과 같다.



[그림 1] 자율 드론의 재난구조 상황과 구조 시나리오

2.1.1 통신 두절 상황

드론이 구조대상자 수색 임무 수행 중 통신 음영 지대에 진입하여 GCS(Ground Control System)와 상호간 통신이 두절된 상황을 가정한다. 이때, 드론이 구조대상자를 식별했을 경우, 기존의 드론은 해당 위치만을 기억하고 지상통제소 또는 기존 계획된 통신 가능한 경로로 이동하고자 하므로 구조대상자에 대한 대응이 제한되지만 제안하는 드론은 다음과 같은 절차로 구조 임무를 효과적으로 수행한다.

- ① 임무 시작: 구조대상자 탐색을 위해 특정 구역을 수색하도록 음성 또는 텍스트를 통해 지시를 받고 임무 시작
- ② 통신 두절 및 구조대상자 발견: 수색 중 통신 음영 지대에서 구조대상자 발견
- ③ 음성 기반 신원 및 상태 확인: 드론은 내장 스피커를 통해 구조대상자의 신원과 부상 정도 등을 확인하는 음성을 출력(예: “저는 구조용 드론입니다. 이름이 무엇인가요? 어디가 아프신가요?” 등)
- ④ 자연어 인식 및 질의: 구조대상자가 신원 및 부상 정도를 대답하고(예: “저는 홍길동입니다. 오른팔이 골절된 것 같고, 복부

쪽에 출혈이 있어요”), 추가 정보를 공유(예: 우측으로 5m 정도에 있는 아내도 구해주세요.”)하면, 드론이 그에 대응하여 추가로 필요한 사항에 대해 질문(예: “출혈이 있다고 하셨는데 긴급 수혈 준비를 위해 혈액형을 알려주세요. 그리고 추가로 구조가 필요한 사람은 없나요?”). 이러한 상호 문답 과정은 효과적인 구조 임무 수행을 위해 자율 드론이 음성인식 및 SLM 기술을 활용하여 필요한 만큼 수행

- ⑤ 분석 및 판단: 다른 구조대상자의 위치와 상태를 추가로 확인 후 수집된 정보를 분석 및 판단
- ⑥ 통신 가능지역 이동 및 전송: 통신 가능지역으로 이동하여 의료기관, 소방본부, 구조 지휘본부 등 관련 기관에 판단결과 전송. 예를 들어, 의료기관에 구조대상자의 부상 정보(신원정보, 부상 상태, 혈액형 등)를 제공하여 긴급 진료 준비를 할 수 있도록 함
- ⑦ 구조대상자 안정 유도: 현장에 복귀하여 구조대상자에게 심리적 안정을 유도할 수 있도록 음성을 통한 안내(예: “신고가 접수되었습니다. 가능한 한 움직이지 말고 조금만 기다려주세요.”)

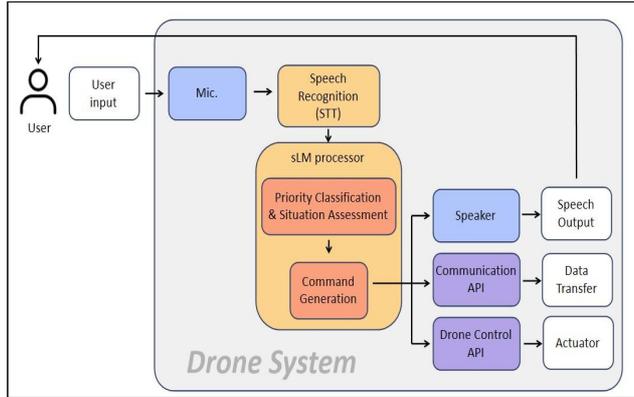
2.1.2 대량 사상자 발생에 따른 다수 드론 임무 수행 상황

다중 차량 추돌, 폭발, 대형 사고 등 수십 명 이상의 사상자가 다수 발생한 심각한 상황이다. 조종사 1명이 드론 1대를 조종하는 기존 시스템에서는 각 조종사별 획득한 정보를 현장지휘자가 취합하여 판단함으로써 실시간 대응이 제한되지만 제안하는 드론은 다음과 같은 절차로 임무를 효과적으로 수행한다.

- ① 임무 시작: 구조대상자 탐색을 위해 특정 구역을 수색하도록 음성 또는 텍스트를 통해 지시를 받고 임무 시작
- ② 구조대상자 발견: 사고현장에서 다수의 구조대상자 발견
- ③ 음성 기반 신원 및 상태 확인: 드론은 내장 스피커를 통해 구조대상자의 신원과 부상 정도 등을 확인하는 음성을 출력(예: “저는 구조용 드론입니다. 이름이 무엇인가요? 어디가 아프신가요?” 등)
- ④ 자연어 인식 및 질의: 구조대상자가 신원 및 부상 정도를 대답하고, 드론은 여러 사상자의 부상 정도(중상/경상/의식불명 등) 등을 판단하기 위한 병렬 질의 수행
- ⑤ 분석, 판단 및 전송: 수집된 정보를 분석 및 판단 후, 관련 기관에 판단결과 전송(예: 1번 드론 수색 진행: 20% / 흰색 카니발(000가1234) 중증 부상자 3명, 검은색 승용차(000나2345) 의식불명 1명 / 의식불명 구조대상자에 대한 우선 조치 필요)
- ⑥ 구조대상자 안정 유도 및 2차 사고 예방: 구조대상자에게 심리적 안정을 유도하고, 2차 사고 예방을 위한 안내 메시지(대피 등) 송출

2.2 자율 드론 시스템 구성

본 연구에서 제안하는 자율 드론 시스템 구성은 [그림2]와 같으며, 재난 상황에서 구조요청자의 음성을 기반으로 자율 드론이 실시간으로 판단을 내리고 적절한 조치를 수행할 수 있도록 설계하였다.



[그림 2] 자율 드론 시스템 구성

1) 음성인식 및 자연어 처리 모듈

우선, 구조요청자(User)로부터 입력된 자연어 음성은 드론에 내장된 마이크(Mic)를 통해 수집된다. 수집된 음성 데이터는 STT(Speech To Text) 모듈[11]을 통해 텍스트로 변환되고, 변환된 텍스트는 SLM 기반의 자연어 처리 모듈로 전달된다. SLM 프로세서는 자연어 텍스트의 의미를 분석하여 구조대상자의 의도와 핵심 정보(예: 인적사항, 피해 유형, 부상범위와 정도 등)를 파악한다.

2) 상황평가 및 우선 임무순위 분류 모듈

다음으로, 상황평가 및 우선순위 분류 모듈에서는 파악된 정보를 기반으로 구조의 긴급성, 환경 위험도 등을 종합적으로 분석하여 구조 임무의 우선순위를 판단한다. 판단결과는 명령 생성 모듈로 전달되어, 드론이 수행할 구체적인 행위명령으로 변환된다. 이때, 사용자에게 추가 정보수집이나 음성안내가 필요할 경우 내장된 스피커를 통해 음성 피드백을 제공하여 양방향 의사소통을 유도할 수 있다. 또한, 판단된 구조 정보는 통신 API(Application Programming Interface)를 통해 의료기관, 소방본부, 구조 지휘본부 등 관련 기관에 전송되며, 제어 API를 통해 드론의 물리적 움직임을 제어하여 비행, 착륙, 물품 투하, 탐색 재개 등을 수행할 수 있다.

이와 같은 구조는 기존의 드론이 단순히 정보를 수집하는 단계 를 넘어, 구조요청자의 음성을 통해 상황을 정확히 이해하고 구조 임무를 능동적이고, 효과적으로 수행함으로써 재난구조 상황에서 구조 임무의 신속성과 효율성을 크게 증가시킬 수 있다.

3. 결론 및 향후 연구 계획

본 연구에서는 재난상황에서 SLM 기반의 자연어 음성인식 처리기술을 활용하여 구조대상자의 음성인식과 상황 파악을 통해 효과적으로 재난구조 임무 수행이 가능한 재난구조용 자율 드론을 제안하였다. 구체적으로, 제안하는 자율 드론의 두 가지 상황별 활용 시나리오를 설명하고, 그러한 임무 수행을 위한 자율 드론 시스템 구성도와 함께 핵심 모듈의 동작 기능을 기술하였다. 이를 통해, 본 연구에서 제안한 자율 드론은 기존의 드론 활용 대비 더욱 효과적인 구조 임무를 수행할 수 있을 것이라 판단한다.

향후 연구은 다음과 같다. 첫째, 재난구조상황에서 구조자의 음성을 통해 구조자의 상황과 필요한 요청사항과 의도를 정확히 이해할 수 있도록 재난구조에 특화된 SLM을 구축할 계획이다. 둘째, 제안 드론의 시스템 구성에 따라 드론을 제작하고 필요한 기술을 식별하여 드론에 설치하여 실제 실험을 통해 효과성을 분석하고 추가로 필요한 성능들을 식별할 계획이다. 셋째, 드론의 소음을 상쇄시켜 구조대상자의 음성을 명확하게 판별하는 기술과 드론이 말하는 소리가 선명하게 들릴 수 있는 기술을 발전시키는데 주안을 두고 연구할 계획이다. 이를 통해 각종 재난상황에서 신속한 대처를 통한 생존율을 향상하는데 기여하고자 한다.

참고문헌

- [1] 임채덕, “첨단 ICT 융합 드론 플랫폼 기술 및 개발 방법론,” 미래융합전략센터 융합연구리뷰, vol. 10, pp. 4-28, 12월, 2024년.
- [2] Plaza, J., “The Commercial Sweet Spot of the Drone Industry,” COMMERCIAL UAV NEWS, July, 2024, <https://www.commercialuavnews.com/the-commercial-sweet-spot-of-the-drone-industry>.
- [3] Cuthbertson, A., “AI model passes Turing Test ‘better than a human’,” Independent, April, 2025, <https://www.independent.co.uk/tech/ai-turing-test-chatgpt-openai-agi-b2728930.html>.
- [4] Strachan, J. W. A., Albergo, D., Borghini, G. et al., “Testing theory of mind in large language models and humans,” Nature Human Behaviour, vol. 8, pp. 1285-1295, July, 2024.
- [5] Caballar, R. D., “What are small language models?,” IBM, October, 2024, <https://www.ibm.com/think/topics/small-language-models>.
- [6] 김근용, 윤기하, 김량수 등, “온디바이스 소형언어모델 기술 개발 동향,” ETRI 전자통신동향분석, vol. 39, no. 4, pp. 82-92, 8월, 2024년.
- [7] Sapien, “The Efficiency, Applications, and Advancements in AI Small Language Models,” May, 2024, <https://www.sapien.io/blog/the-efficiency-applications-and-advancements-in-ai-small-language-models>.
- [8] Papyan, N., Kulhandjian, M., Kulhandjian, H. et al., “AI-Based Drone Assisted Human Rescue in Disaster

Environments: Challenges and Opportunities,” Pattern Recognit and Image Analysis, vol. 34, pp. 169-186, April, 2024.

- [9] Judson, J., “Voice-controlled drones a military game-changer, Primordial Labs says,” DefenseNews, March, 2025, <https://www.defensenews.com/industry/techwatch/2025/03/04/voice-controlled-drones-a-military-game-changer-primordial-labs-says>.
- [10] 정수진, “DJI, 손바닥 크기의 비로그용 드론 ‘DJI 니오’ 출시,” CAD&Graphics, 9월, 2024년, <https://www.cadgraphics.co.kr/newsview.php?pages=news&sub=news02&catecode=3&num=75269>.
- [11] Microsoft, “What is Speech to Text?,” March, 2025, <https://learn.microsoft.com/en-us/azure/ai-services/speech-service/speech-to-text>.