

전장 상황 인지 보고서 생성을 위한 Text-to-Text 멀티 태스크 학습

허종국¹ · 임새린¹ · 안시후¹ · 박진혁¹ · 이영재¹ · 조용원¹ · 목충협¹ · 조준호² · 김성범[†]

¹고려대학교 산업경영공학과 / ²리얼타임비주얼(주) 기술연구소

Text-to-Text Multi-Task Learning to Generate Battlefield Situation Awareness Report

Jongkook Heo¹ · Sae Rin Lim¹ · Sihuh Ahn¹ · Jinhyeok Park¹ · Young Jae Lee¹ · Yongwon Jo¹ · Chunghyup Mok¹
Junho Cho² · Seoung Bum Kim¹

¹Department of Industrial and Management Engineering, Korea University

²R&D Center, REALTIMEVISUAL Inc.

Advances in new communication technologies enable commanders to collect various information in battlefield situations. However, it is difficult to make quick and accurate decisions on the battlefield because of vast amount of information. To address this problem, several studies attempt to change tabular data into an easy-to-understand text format. Existing table-to-text studies are not suitable for battlefield situations because they use specific domain data such as WIKIBIO and WIKITABLETEXT. In this study, we propose a table-to-text transfer transformer (TaT4) that uses special tokens to transform log table data into a single sequence to preserve table information. Moreover, the proposed TaT4 uses multi-task learning that can leverage cross-task data of types in a single model to improve generalization performance. We conduct experiments on eight datasets generated from three Korean defense modeling and simulations (M&S) of battlefield situations in the Army, Air Force, and Navy. The proposed TaT4 outperforms the existing table-to-text models.

Keywords: Battlefield Situation Awareness, Multi-Task Learning, Table-To-Text, Text-To-Text

1. 서론

전장 상황 정보를 적시에 인지하는 것은 지휘관의 상황 판단 및 의사 결정에 영향을 미치기 때문에 매우 중요하다(Kim *et al.*, 2021). 정찰기 및 통신 기술의 고도화에 따라 지휘관은 전장 상황에서 다양한 정보를 수집할 수 있게 되었지만, 복잡한 전장 상황에서 필요한 정보를 추출하고 분석하는 것은 매우 어렵다(Jin *et al.*, 2020; Kim *et al.*, 2021). 최근 우리 군은 입출력을 위한 단순한 질의형 데이터베이스인 전장관리체계

(KJCCS, JFOS-K, ATCIS)를 통해 지휘관의 지휘 결심을 보조하고 있지만 정보 판독관의 해석이 개입해야 하는 한계가 있다(Han *et al.*, 2019).

최근 국방 개혁과 병력 감축으로 인해 국방 분야에서는 인공지능을 적용한 다양한 연구가 진행되고 있다. Moon(2018)은 육군 전투지휘훈련단의 근접 전투 훈련결과 데이터에서 의사결정나무 및 인공 신경망 모델을 적용하여 전투 피해 평가 요소를 예측함으로써 한국군의 독자적 전장 상황 평가 모형을 구축하였다. 또한 국방 모델링 및 시뮬레이션(M&S)의 중요성이

본 연구는 방위사업청과 방위산업기술지원센터의 지원으로 수행되었습니다. (UC200020D)

† 연락처: 김성범 교수, 02841, 서울특별시 성북구 안암로 145 고려대학교 산업경영공학부, Tel : 02-3290-3397, Fax : 02-929-5888,

E-mail : sbkim1@korea.ac.kr

2022년 9월 13일 접수; 2022년 10월 11일 수정본 접수; 2022년 10월 26일 게재 확정.

대두되면서 computer generated force(CGF) 행동 규칙을 모의하는 연구 또한 급증하고 있다. Han *et al.*(2011)은 CGF 행동을 규칙 기반으로 모의하는 과정에서 퍼지 추론을 적용하여 이진 값으로 표현되지 않는 주관적 군사 모의 논리를 가능하게 하였다. Lee *et al.*(2012)은 CGF의 행동을 규칙 기반이 아닌 partially observable Markov decision process(POMDP)로 모델링함으로써 복잡한 전장 상황에서 전문가의 도움 없이 최적의 행동 규칙을 모의할 수 있음을 보여주었다. 나아가 Toghiani-Rizi *et al.*(2017)은 Mnih *et al.*(2013)가 제안한 deep Q-Network(DQN)과 Mnih *et al.*(2016)가 제안한 asynchronous advantage actor critic (A3C) 강화학습 알고리즘을 적용하여 CGF의 행동 규칙을 모델링 하였다.

인공지능을 적용한 전장 상황 인식에 대한 연구는 텍스트로 이루어진 전장 상황 정보로부터 지식 베이스를 구축하여 정보를 추출하는 연구가 주로 진행되었다. Jo *et al.*(2020)은 Devlin *et al.*(2018)이 제안한 bidirectional encoder representations from transformers(BERT)를 활용하여 군사 정보를 담고 있는 텍스트에서 개체명을 인식한 후, 개별 정보를 ‘주어(노드)-관계(엣지)-목적어(노드)’ 형태의 트리플렛 형태로 저장하는 지식 베이스를 구축하는 방안을 제시하였다. Jin *et al.*(2020)은 지휘관이 지식 베이스에 입력한 질의로부터 하이퍼 그래프를 통해 연관된 질의를 추출함으로써 전장 상황에 대한 복합적이고 다양한 관점의 추론이 생성될 수 있음을 실험적으로 증명하였다. 또한 Kim *et al.*(2021)은 지식 베이스 내에 존재하는 개별 정보를 역으로 문장화한 후 임베딩 모델을 활용하여 유사도 기반 관심 정보를 추출할 수 있음을 보여주었다.

지능형 전장 상황 인식을 위한 선행 연구들은 텍스트로 이루어진 전장 상황 데이터에서 개체 관계를 추출하여 그래프 형태의 지식 베이스로 저장함으로써 유사한 정보를 추출하거나 복합적인 추론이 가능함을 보여주었다. 하지만 정찰 및 탐지 기기로부터 생성되는 원시 데이터는 수치형 로그 데이터이기 때문에, 지식 베이스 구축에 앞서 해당 데이터로부터 전장 상황을 올바르게 묘사하는 인지 보고서를 생성하는 것이 우선시 되어야한다. 하지만 훈련장 크기나 훈련 비용 등의 물리적, 비용적 문제와 안보 문제로 인해 실제 전장 상황에서 정찰 기기로부터 전장 상황 로그 데이터를 확보하는 것은 어렵다(Kim *et al.*, 2019). 따라서 본 연구에서는 3개의 국방 M&S인 CJ21(창조21), CG(창공), CH(청해) 모델을 활용하여 다양한 시나리오에서 전장 상황 로그 데이터를 추출하고, 해당 데이터로부터 다양한 유형의 전장 상황 인지 보고서를 생성하는 Table-to-Text 연구를 진행하였다.

기존 Table-to-Text 연구는 다수의 필드(field)와 값(value)으로 이루어진 단일 테이블에서 보고서를 생성하기 위해 테이블 데이터에 적합한 전처리 방식, 어텐션 메커니즘, 모델 구조를 제안하는 방식으로 발전되어왔다. Lebreton *et al.*(2016)은 값의 단어를 필드명과 단어뿐만 아니라 필드 내 위치 임베딩에 대한 결합으로 정의하는 전처리 방식을 제안하였다. Liu *et al.*(2018)

은 테이블 구조를 효율적으로 인코딩하기 위해 새로운 필드 게이트 인코더 구조와 듀얼 어텐션 메커니즘을 제안하였다. Bao *et al.*(2018)은 copy mechanism(Gu *et al.*, 2016; Nallapati *et al.*, 2016)을 통해 디코더가 지명이나 인물의 이름 등 말뭉치에 없는 고유 명사를 생성할 수 있게 하였다. 선행 연구들은 하나의 정보를 통해 보고서를 생성하는 단일 정보 테이블에만 적용 가능하다는 한계가 있다. 예를 들어, table-to-text 연구에서 벤치마크 데이터로 많이 사용되는 WIKIBIO는 테이블 정보를 통해 한 인물의 전기를 생성하지만 여러 인물에 대한 정보가 동시에 입력되는 상황은 고려하지 않는다. 하지만 전장 상황 인지 보고서는 여러 장소와 시간을 통해 수집된 정찰 및 첩보를 바탕으로 정보가 생성되는 경우가 대부분이다. 또한 전장 상황 인지 보고서는 부대의 종류나 상황에 따라 다양한 유형이 존재하기 때문에 보고서별로 모델을 구축할 경우 시간과 데이터 수집 비용이 많이 요구된다. 이러한 문제들로 인해 기존 선행 연구를 전장 상황 인지 보고서에 바로 적용하는 것은 어려운 실정이다.

본 연구에서는 트랜스포머를 기반으로 다중 정보를 처리하기 위한 전처리 기법과 멀티 태스크 학습을 접목하여 다양한 종류의 전장 상황 인지 보고서를 하나의 모델로 생성할 수 있는 table-to-text transfer transformer(TaT4)를 제안한다. TaT4는 테이블 구조를 모델에 효율적으로 입력하기 위하여 스페셜 토큰을 활용한다. 스페셜 토큰이 테이블 필드와 그 필드에 해당하는 정보를 구분할 수 있게 해줌으로써 모델이 테이블 구조에 대한 정보를 전달받는다. 제안 전처리 기법은 하나의 필드에 여러 정보가 있는 다중 정보 테이블로도 쉽게 확장이 가능하다. 또한 TaT4는 보고서 유형을 태스크 토큰으로 함께 입력하여 여러 보고서를 하나의 모델로 생성할 수 있도록 멀티 태스크 학습을 수행한다. 모델은 보고서 유형을 통해서 입력된 테이블 정보로부터 어떤 보고서를 생성해야 하는지 학습할 수 있게 되며, 최종적으로 하나의 모델을 통해 여러 보고서를 효율적으로 학습할 수 있게 된다. 또한 멀티 태스크 학습은 개별 데이터가 부족하더라도 연관된 데이터를 통합하여 단일 모델로 학습하는 방법론으로, 데이터 증가 효과와 개별 태스크의 과적합을 방지하는 효과가 있다(Liu *et al.*, 2019). TaT4 역시 다양한 보고서 데이터를 통합하여 학습에 활용하기 때문에 상대적으로 데이터 수가 적은 보고서라도 효과적으로 학습할 수 있다. 본 연구에서는 TaT4의 한국어 생성 능력을 향상시키기 위해 한국어 3억 7천만 문장과 한/영 뉴스 데이터 등을 통해 사전 학습된 KE-T5(Shin *et al.*, 2021)로 모델을 초기화한다.

본 연구에서는 총 8개의 보고서 생성 작업을 통해서 TaT4의 성능을 평가하였다. 기존 table-to-text 방법론과 비교를 통해서 제안 방법론의 우수성을 입증하였으며, 다양한 비교 실험을 통해서 TaT4의 효용성을 보여주었다. 또한 셀프 어텐션 분석을 통해서 제안 방법론이 보고 부대와 타겟 부대의 상황 변화에 따라 중요한 전장 정보를 유동적으로 선별할 수 있다는 것을 보여주었다. 본 논문의 주요 기여점은 다음과 같다.

- 국방 M&S를 활용해 전장 상황 로그 데이터를 수집하고 이를 통해 전장 상황 인지 보고서를 생성하는 연구를 수행하였다. 따라서 본 연구는 향후 실제 전장 상황에서 정찰 및 탐지 기기로부터 수집된 수치형 로그 데이터에서 전문가의 개입없이 필요한 정보를 제공하여 지휘관의 의사 결정을 보조할 수 있을 뿐만 아니라 지능형 전장 상황 인식 서비스 개발을 위한 비정형 데이터를 생성할 수 있다.
- 필드 및 값의 영역을 구분하는 스페셜 토큰을 활용함으로써 table-to-text 문제를 text-to-text 문제로 정의하였다. 제안 전처리 기법은 단일 정보 테이블만을 고려한 기존 연구와 달리 다중 정보 테이블에서도 적용될 수 있다.
- 제안 전처리 기법을 통해 보고서 유형을 태스크 토큰으로 함께 입력하여 멀티 태스크 학습을 수행함으로써 하나의 모델을 통해 효율적으로 다양한 종류의 전장 상황 인지 보고서를 생성할 수 있다. 이는 특정 보고서에 과적합되지 않도록 하는 규제 효과와 데이터 증가 효과로 개별 보고서 개수가 부족하더라도 강건한 성능을 나타내는 모델을 구축할 수 있다.

본 논문의 구성은 다음과 같다. 제2장에서는 국방 M&S에서 사전에 정의한 시나리오를 통해 전장 상황 로그 데이터를 수집하는 과정을 소개하고 데이터에 대해 설명한다. 제3장에서는 테이블을 텍스트로 변환하는 전처리 기법과 태스크 토큰을 활용한 멀티 태스크 학습 과정을 서술한다. 제4장에서는 실험 결과를 통해 제안하는 전처리 기법 및 멀티 태스크 학습의 우수성을 증명한다. 또한 어텐션 스코어 맵을 통해 보고 부대와 타겟 부대의 상황 변화에 따라 중요한 전장 정보를 선별할 수 있음을 서술한다. 제5장에서는 결론 및 향후 실제 전장 상황에서의 기대 효과를 서술한다.

2. 국방 M&S를 활용한 전장 상황 로그 데이터 생성

본 연구에서는 3개의 국방 M&S인 창조21, 창공, 청해 모델을 통해 학습을 위한 전장 상황 로그 데이터와 레이블을 생성하였다. 창조21, 창공, 청해 모델은 실제 국군의 군사 작전 모의 및 한미 연합 훈련에 활용되고 있는 위게임 모델이며, 각각의 모델은 위게임 개발 사업에서 시범평가 절차를 거쳐 검증되었다. 구체적으로 창조21은 육군 전투지휘훈련단, 창공은 공군 작전사령부, 청해는 해군전력 분석시험평가단에서 기술 이전을 받아 개발되었다. 창조21 모델은 육군의 사단/군단급 전투 지휘를 위한 지상전 모의 모델이며 보병 혹은 전차로 이루어진 가상군 개체를 통해 대대 규모 임무를 수행한다. 창공 모델은 공대공, 공대지, 방공포 및 방공 통제 등의 다양한 공군 편대 임무를 수행하는 공중전 모의 모델이다. 청해 모델은 해군 작전사령부 예하의 함대급 지휘관과 참모의 훈련을 지원하는 해상전 모의 모델이며 함정, 잠수함, 기뢰, 어뢰를 개체로 사용한다. 본 연구에서는 보안상의 이유로 국방 M&S 내의 개체명, 지역의 좌표를 수정하여 시뮬레이션을 진행하였음을 밝힌다.

<Figure 1>은 국방 M&S에 시나리오를 입력하여 전장 상황 로그 데이터를 생성하는 전체 프로세스를 나타내며, <Figure 2>는 전장 상황 모의 시나리오의 예시이다. 각 M&S는 사전에 정의된 시나리오를 입력 받아 가상군 개체 간 상호 작용을 통해 반복 모의를 수행한다. M&S의 가상군 개체는 M&S에 내재된 행동 규칙에 따라 이동 및 전투를 수행한다. 반복 모의를 통해 수집된 초기 로그 데이터는 부대/개체 속성, 교전 인터렉션, 지형/지물/환경 속성 등으로 구성되어 있다. 보고서 종류에 따라 사용 가능한 로그 정보와 보고서 양식이 상이하기 때문에 보고서 유형에 따른 정형 전장 자료 생성기를 통해 학습을 위한 데이터로 변환 후 데이터베이스에 저장된다.

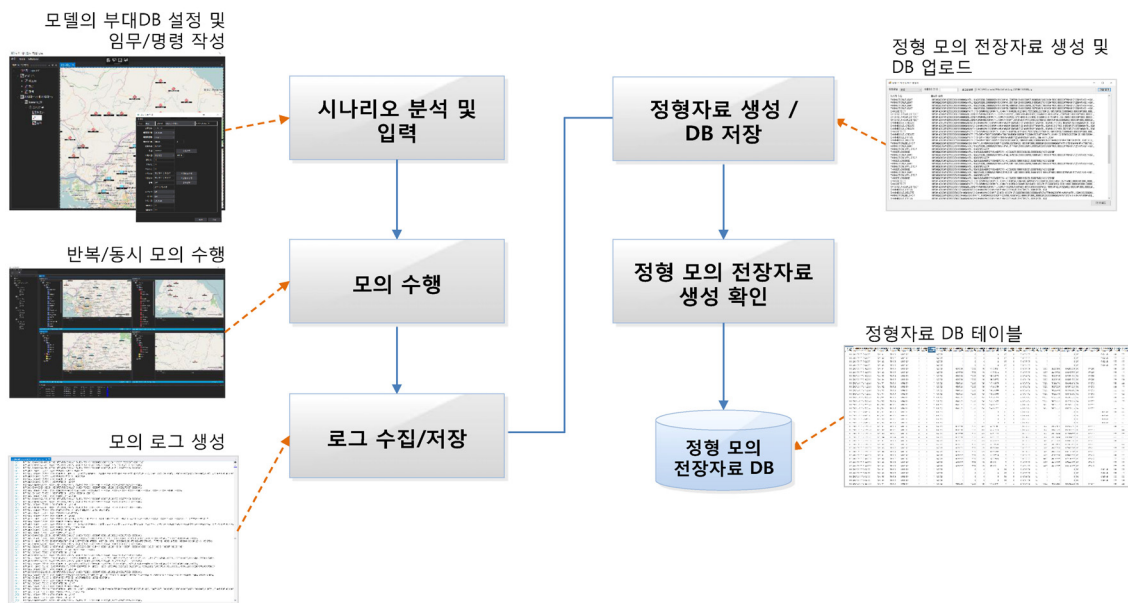


Figure 1. Creating Process of Battlefield Log Data Through Korean Defense M&S

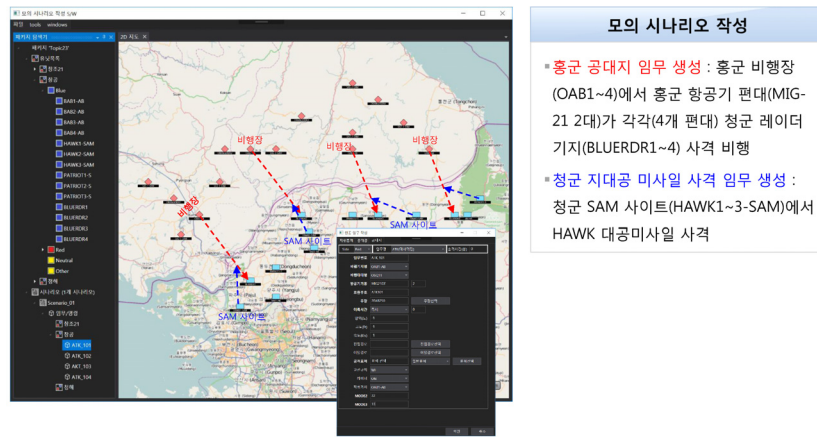


Figure 2. An Example of Battlefield Simulation Scenario

정형 전장 자료 생성기는 M&S의 로그 데이터에서 보고서 유형별 입력값 테이블과 타겟 레이블을 생성한다. 타겟 레이블은 보고서별로 사전에 정의한 규칙에 따라 생성되며, 보고 내용의 순서 및 어미 등에 일부 무작위성을 부여하여 다양한 문장을 생성한다. 또한 CGF 개체의 상태와 적/아 정보의 탐지 여부에 따라 보고 내용의 차이가 존재한다. <Figure 3>은 창조

21 모델을 통해 생성한 교전 결과 보고서(engagement result)의 타겟 레이블의 예시이며, 적/아 정보 탐지 여부에 따라 보고 내용이 상이함을 알 수 있다. X는 해당 정보가 없음을 의미한다. 각 M&S별로 사용한 시나리오 개수와 생성한 보고서 종류 및 그에 따른 학습 데이터 개수는 <Table 1>과 같다. 단일 로그에서 단일 보고서를 생성하는 나머지 7개의 보고서와 달리 창

시간	2019년 04월 15일 10시 31분
보고서 종류	적 접촉 상황보고
전투피해 종류	X
보고부대	A연대1중대
보고부대 지역	강릉지역
보고부대 위경도	37.791629N, 128.911409E
보고부대 피해량	X
...	...
적군 부대 종류	경보병
적군 부대 지역	강릉지역
적군부대 위경도	37.791628N, 128.910273E
적군 부대 활동	남서 이동
적군 부대 규모	100

2019년 04월 05일 10시 31분 강릉지역(37.791629N, 128.911409E)에 위치한 아군 A연대1중대는 적군 경보병부대와 접촉함. 관측된 적은 강릉지역(37.791628N, 128.910273E)에서 남서쪽으로 이동중이며, 적 규모는 100여명임

시간	2019년 04월 05일 13시 29분
보고서 종류	전투 피해 상황 보고
전투피해 종류	적 포병 사격
보고부대	B사단 지휘소
보고부대 지역	X
보고부대 위경도	X
보고부대 피해량	2명 사망/0명 부상/장비 3점 완파/0점 반파
...	...
적군 부대 종류	X
적군 부대 지역	X
적군부대 위경도	X
적군 부대 활동	X
적군 부대 규모	X

2019년 04월 05일 13시 29분 아군 B사단지휘소가 적 포병사격에 의해 2명이 사망하였으며, 장비 3점이 완파 되었습니다.

Figure 3. Examples of Engagement Result Reports from CJ21

Table 1. Description of Data Generated from Three Korean M&S

M&S	Dataset	Number of Scenarios	Number of Data	Data Type
CG	Anti-Aircraft Tracking	2	59	Multi-Row
	Anti-Aircraft Radar Detection	2	267	Single-Row
CH	Warship Observation	9	6995	Single-Row
	Engagement Situation	1	75	Single-Row
CJ21	Engagement Result	5	111	Single-Row
	Ground Forces Observation	6	318	Single-Row
	Legion UAV Video	19	443	Single-Row
	Reconnaissance Wing Video	14	3418	Single-Row

공 모델에서 생성한 대공 항적 보고서(anti-aircraft tracking)는 아군 레이더 기지에서 여러 시점 동안 적 항공기 편대의 위치를 기록한 후 하나의 보고서를 생성하는 다중 로그 단일 보고서 데이터이며 각 보고서별 사용되는 행의 개수가 가변적이다. 본 연구에서는 제안 방법론이 단일 정보 테이블 뿐만 아니라 다중 정보 테이블에서도 적용이 가능함을 증명한다.

3. 제안 방법론

본 장에서는 제안 방법론 TaT4에 대해 서술한다. TaT4는 두 단계로 나뉘며, 첫 번째는 text-to-text 학습을 수행하기 위해 테이블 데이터를 하나의 텍스트로 변환하는 전처리 기법이다. 본 연구는 필드명과 값을 구분하는 스페셜 토큰을 통해 테이블 구조를 표현한다. 두 번째는 태스크 토큰을 활용한 멀티 태스크 학습이다. 본 연구에서는 연관된 데이터를 하나로 통합하여 멀티 태스크 학습을 수행하며, 생성하고자 하는 보고서 종류를 태스크 토큰으로 텍스트 앞부분에 추가한다. 본 제안 방법론의 최종 프레임워크는 <Figure 4>와 같다.

3.1 스페셜 토큰을 활용한 Table-to-Text 전처리

Text-to-text 학습을 수행하기 위해서는 입력 값인 테이블 데이터를 하나의 텍스트 데이터로 변환해야 한다. 테이블 데이

터를 텍스트로 변환하는 가장 단순한 방법은 필드명과 필드에 해당하는 값을 단순히 이어 붙이는 것이다. 하지만 이러한 전처리는 필드명과 해당 값이 구분되지 않아 테이블의 구조적 특징을 반영할 수 없다(Liu *et al.*, 2018). 기존 table-to-text 연구들은 단일 정보 테이블의 구조적 특징을 반영하기 위해 필드명과 필드값 위치 정보를 추가로 제공할 경우 성능이 향상됨을 실험적으로 증명하였다(Lebret *et al.*, 2016; Liu *et al.*, 2018). 따라서 본 연구에서는 스페셜 토큰을 활용하여 테이블 데이터의 구조적 특징을 반영할 수 있는 전처리 기법을 제안한다. 각 필드에 해당하는 값은 단일 값인 경우도 있으며 여러 행으로 이루어진 다중 값일 수도 있다. 또한 각 필드명과 필드값은 하나 이상의 토큰으로 이루어질 수 있다. 본 연구에서는 필드명에 대한 스페셜 토큰 <c>와 </c>를 각각 필드명의 앞 뒤에 추가함으로써, 하나 이상의 토큰으로 이루어진 필드명 영역을 구분하였다. 필드값 또한 스페셜 토큰 <v>와 </v>를 앞과 뒤에 추가하여 필드값 영역을 표시하였다. 예를 들어 스페셜 토큰을 활용하지 않는 경우 <Figure 5>의 단일 정보 테이블 Unit ID는 [Unit, ID, 1001]의 시퀀스로 표현할 수 있지만, 첫 번째 필드명이 Unit인지 Unit ID인지 구분하기 어렵다. 하지만 스페셜 토큰을 활용할 경우 <Figure 5>의 Unit ID는 [<c>, Unit, ID, </c>, <v>, 1001, </v>]로 표현할 수 있기 때문에 여러 토큰으로 이루어진 필드명과 필드값의 범위를 구분하기 쉽다.

기존 table-to-text 연구들은 단일 정보 테이블에만 적용이 가능하며 두 개 이상의 정보를 가진 다중 정보 테이블에는 적용

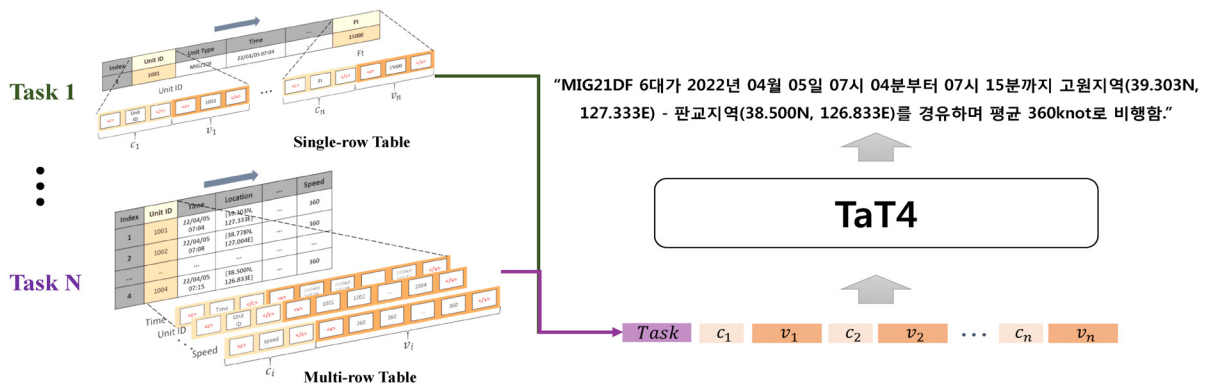


Figure 4. Overall Framework of the Proposed TaT4

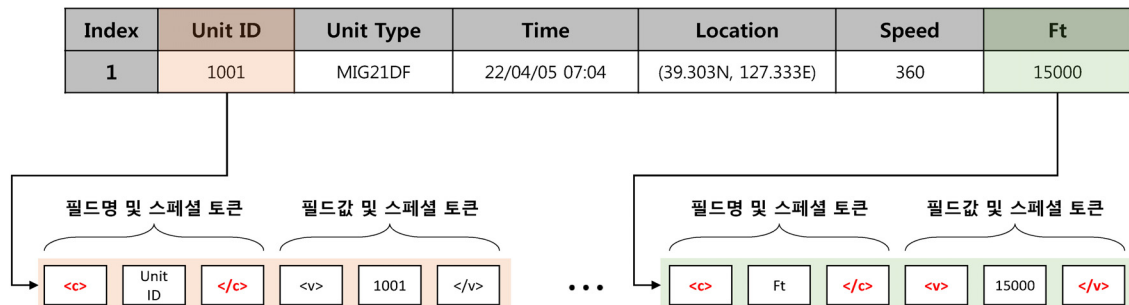


Figure 5. The Preprocessing Process for Single-Row Table

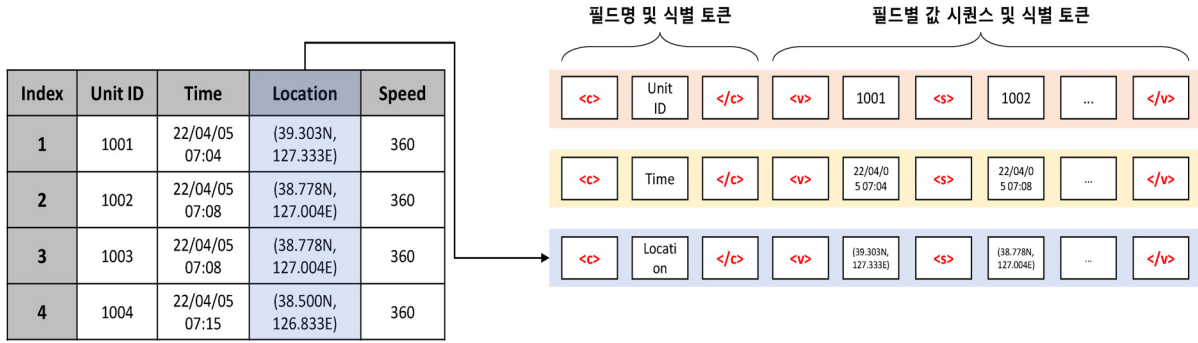


Figure 6. The Preprocessing Process for Multi-Row Table

할 수 없다. 하지만 제안 전처리 기법은 한 필드 내 여러 값들 사이에 이를 구분해주는 스페셜 토큰을 추가함으로써 다중 정보 테이블에도 쉽게 적용할 수 있다. 예를 들어 <Figure 6>의 Location 필드값의 시퀀스가 <v>, 39., 303, N, 127.,333, E, 38., 778, N,...</v>로 표현 될 경우, 같은 필드 내의 다른 값을 구분하기 어렵다. 하지만 각 필드값의 영역을 구분하기 위한 스페셜 토큰 <s>를 추가하면 <Figure 6>의 오른쪽과 같이 다중 정보 테이블을 간단하게 표현할 수 있다. 따라서 제안 전처리 기법을 활용한 방법론은 보고서 생성시 어떤 필드가 중요한지 필드별 어텐션 스코어를 세부적으로 산출할 수 있다.

3.2 모델 아키텍처

본 연구에서 제안하는 TaT4는 자연어 처리에서 뛰어난 성능을 보이고 있는 트랜스포머(Vaswani *et al.*, 2017)를 활용한 인코더와 디코더로 구성되어 있다. 3.1장에서 설명한 전처리 과정을 거친 테이블 데이터는 길이가 n 인 시퀀스로 변환된다. 각 시퀀스 값들은 d 차원으로 임베딩되어 입력 시퀀스 $X=(x_1, x_2, \dots, x_n) \in R^{n \times d}$ 로 토큰화 된다. 인코더는 입력 시퀀스 X 를 입력 받아 표현 벡터 $Z=(z_1, z_2, \dots, z_n) \in R^{n \times d}$ 을 산출한다. 디코더는 표현 벡터 Z 와 이전 시점 $t-1$ 까지 생성된 보고서의 내용인 시퀀스 $\hat{y}_{1:t-1}$ 가 주어졌을 때, 다음 시점의 토큰 \hat{y}_t 을 생성한다. 모델의 목적 함수는 식 (1)과 같이 모델의 파라미터 θ 에 대하여 길이가 n' 이고 각 토큰의 임베딩 차원이 d 인 정답 보고서의 시퀀스 $Y=(y_1, y_2, \dots, y_{n'}) \in R^{n' \times d}$ 가 생성될 로그 우도를 최대화하는 것이다.

$$\begin{aligned} \text{Max}_{\theta} \log p(y_1, y_2, \dots, y_{n'} | x_1, x_2, \dots, x_n, \theta) \\ = \text{Max}_{\theta} \sum_{t=1}^{n'} \log p(y_t | Z, y_1, \dots, y_{t-1}, \theta) \end{aligned} \quad (1)$$

인코더와 디코더는 각각 멀티 헤드 셀프 어텐션 레이어와 다중 레이어 퍼셉트론으로 구성된 N 개의 블록으로 이루어져 있다. 셀프 어텐션은 테이블 시퀀스를 이루는 각 토큰의 표현 벡터를 효율적으로 인코딩하기 위해 분석 대상이 되는 토큰과

나머지 토큰에 대한 가중치를 계산하는 메커니즘이다. 셀프 어텐션을 통한 토큰의 가중합 벡터는 수식 (2)와 같이 각각 분석 대상이 되는 단어를 표현하는 쿼리(query; q), 쿼리와 유사도를 계산하기 위한 키(key; k), 키 값에 해당하는 토큰 벡터인 밸류(value; v)를 통해 구한다. q, k, v 는 식 (3)과 같이 d_{model} 차원을 가진 토큰 임베딩 벡터인 Q, K, V 에 대하여 각각 linear projection 행렬인 W_i^Q, W_i^K, W_i^V 를 곱하여 구해지며, q 와 k 의 유사도는 내적을 통해 계산된다. 이후 q 와 k 의 차원 증가에 따른 발산을 방지하기 위해 k 의 차원인 d_k 의 제곱근 값으로 스케일링한 후, 소프트맥스(softmax) 함수를 통해 0에서 1사이의 값으로 어텐션 스코어가 산출되어 v 의 가중치로 활용된다. 또한 식 (4)과 같이 하나의 헤드가 아닌 서로 다른 h 개의 헤드로부터 어텐션 스코어를 산출하여 합한 후, linear projection 행렬인 W_i^O 를 통해 테이블의 시퀀스를 구성하는 개별 토큰에 대한 최종적인 표현 벡터를 산출한다.

$$\text{Attention}(q, k, v) = \text{softmax}\left(\frac{qk^T}{\sqrt{d_k}}\right)v \quad (2)$$

$$q = QW_i^Q, k = KW_i^K, v = VW_i^V \quad (3)$$

$$\text{where } W_i^Q, W_i^K, W_i^V \in R^{d_{model} \times d_k}, W_i^O \in R^{(h \times d_k) \times d_{model}}$$

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h)W_i^O \quad (4)$$

$$\text{where } \text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V)$$

3.3 전장 상황 인지 보고서 생성을 위한 Text-to-Text 멀티 태스크 학습

전장 상황 인지 보고서는 탐지하고자 하는 개체의 종류와 보고 형식에 따라 다양한 종류가 있다. 예를 들어 지대공 상황에서 적 전투기의 비행 항적을 보고하는 상황이 있을 수 있고, 적/아군 지상 부대의 근접 전투 상황에서 아군 부대의 피해 규모를 보고하는 상황이 존재할 수도 있다. 하지만 모든 보고 상황에 따라 보고서 생성 모델을 학습하는 것은 많은 시간이 요구되기 때문에 현실적으로 적용하기가 어렵다. 이 문제를 해결하기 위해 TaT4는 Raffel *et al.*(2019)에서 활용된 태스크 토

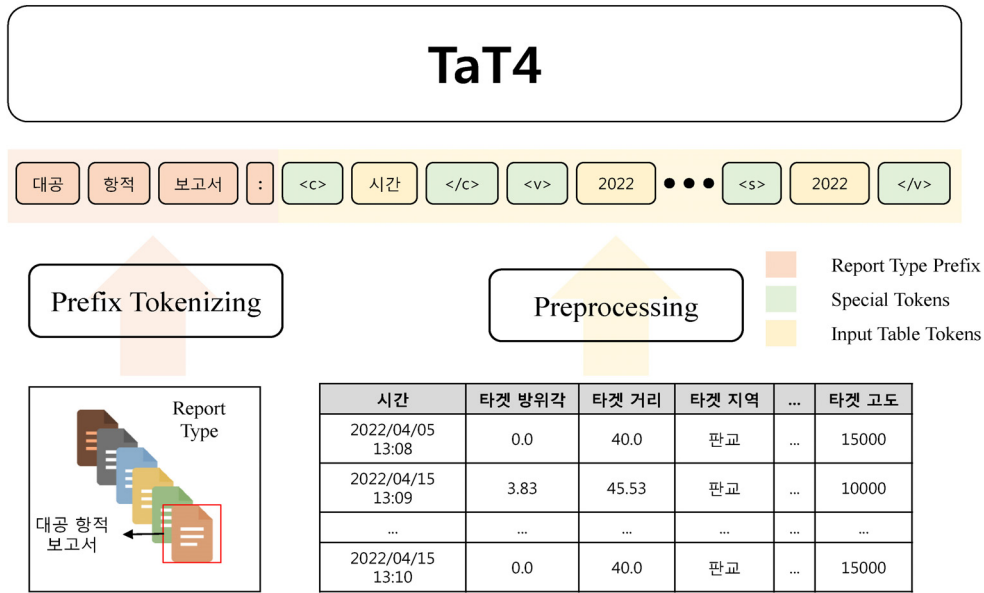


Figure 7. Multi-Task Learning with Report Type Prefix

큰을 다중 보고서 생성으로 확장한다. TaT4는 보고서 유형을 토큰화하여 모델에 추가적으로 입력해줌으로써 간단하고 효과적인 멀티 태스크 학습을 수행한다. <Figure 7>은 TaT4의 멀티 태스크 학습을 보여준다. <Figure 7>에서 볼 수 있듯이 TaT4는 보고서 유형을 토큰화하여 전처리된 테이블 시퀀스 앞에 추가한 후 모델에 입력한다. 모델은 학습 과정을 거치면서 보고서 유형 정보와 테이블 정보 간의 관계를 학습하여 적절한 보고서를 생성하게 된다. 또한 멀티 태스크 학습은 여러 보고서를 하나의 데이터셋으로 통합하여 단일 모델로 학습하기 때문에 학습 데이터가 증가하는 효과가 있으며, 개별 보고서 생성 태스크에 대한 과적합을 방지할 수 있다(Liu *et al.*, 2019).

4. 실험 결과

본 장에서는 3개의 국방 M&S인 창조21, 창공, 청해 모델을 통해 얻은 전장 상황 로그 데이터를 통해 제안 방법론을 학습시키고 모델의 성능을 평가하였다. 구체적인 실험 설정 및 평가 지표에 대한 설명과 함께 다양한 방법론과의 비교 실험을 통해서 제안 방법론의 우수성을 입증하였다. 또한 어텐션 스코어 시각화를 통해서 제안 방법론이 각 보고서마다 중요한 필드를 식별하고 이를 반영하여 보고서를 생성할 수 있다는 것을 입증하였다.

4.1 실험 설정 및 평가지표

본 연구에서 제안하는 TaT4는 트랜스포머(Vaswani *et al.*, 2017)와 동일한 구조를 가지고 있으며, 옵티마이저는 Adafactor

(Shazeer *et al.*, 2018)를 사용하였다. 학습률은 0.001, 배치 사이즈는 64, 에포크는 50으로 3회 반복 실험을 진행하였다. 특히 본 연구에서는 TaT4의 한국어 생성 능력을 향상시키기 위해 한국어 3억 7천만 개 문장과 한/영 뉴스 데이터에서 구축한 총 64,000개 단어를 통해 사전 학습된 KE-T5(Shin *et al.*, 2021)로 모델 파라미터를 초기화하였다. 또한 TaT4는 3개의 국방 M&S에 등장한 지명, 무기 및 개체명과 같은 고유명사들과 테이블 구조를 표현하는 스페셜 토큰들을 포함해 기존 사전에 존재하지 않는 395개의 단어를 추가하여 총 64,395개로 학습을 수행하였다. 테스트 단계에서는 단어 생성 시 빔 서치(beam search)를 사용하였으며 빔의 개수는 3으로 설정하였다. 각 전장 상황 인지 보고서별로 학습, 검증, 테스트 데이터의 비율은 7:1:2로 설정하였다. 멀티 태스크 학습에는 각 보고서별 학습 데이터만을 사용하여 학습하였고, 보고서별로 검증 데이터에 대해 최적의 검증 성능을 보여주는 모델에 대해 각각 테스트를 진행하였다(Raffel *et al.*, 2019).

평가지표는 기존 table-to-text 연구에서 많이 사용되는 ROUGE-4 F1(Lin, 2004)과 BLEU-4(Papineni *et al.*, 2002)를 활용하였다(Lebret *et al.*, 2016; Liu *et al.*, 2018). ROUGE는 식 (5), (6), (7)과 같이 정답 문장과 예측 문장의 n-gram에 대한 점수로 recall과 precision, F1 점수를 계산할 수 있다. 식 (5)와 식 (6)의 references는 정답 문장의 집합, 식 (6)의 candidates는 모델이 생성한 예측 문장의 집합을 의미하며 본 연구에서는 복수의 정답과 예측을 허용하지 않고 단일 정답 보고서에 대한 하나의 예측 값에 대해 점수를 계산하였다. $Count(gram_n)$ 는 정답 혹은 예측 문장의 n-gram의 개수를, $Count_{match}(gram_n)$ 는 정답 문장과 예측 문장에 공통적으로 있는 n-gram의 개수를 나타낸다. 식 (5)는 정답 문장의 n-gram개수에 대해 예측 문장과 실제 문장이 겹치는 n-gram의 개수, 식 (6)은 예측 문장의 n-gram개수에 대해 예측 문장과 실제 문장이 겹치는 n-gram의

개수로 계산되며, 식 (7)은 recall과 precision의 조화 평균으로 계산된다.

$$ROUGE-N_{precision} = \frac{\sum_{S \in \{References\}} \sum_{gram_n \in S} Count_{match}(gram_n)}{\sum_{S \in \{References\}} \sum_{gram_n \in S} Count(gram_n)} \quad (5)$$

$$ROUGE-N_{precision} = \frac{\sum_{S \in \{References\}} \sum_{gram_n \in S} Count_{match}(gram_n)}{\sum_{S \in \{Candidates\}} \sum_{gram_n \in S} Count(gram_n)} \quad (6)$$

$$ROUGE-N_{F1} = 2 \times \frac{ROUGE-N_{precision} \times ROUGE-N_{recall}}{ROUGE-N_{precision} + ROUGE-N_{recall}} \quad (7)$$

BLEU는 식 (6)과 같은 n-gram precision 기반 점수이다. 하지만 BLEU는 생성 문장이 과도하게 길어지거나 같은 단어를 반복적으로 생성할 경우 페널티를 부과한다. 식 (8)과 같이 n-gram precision을 계산할 때 $Count_{match}(gram_n)$ 대신 $Count_{dip}(gram_n)$ 을 사용한다. 식 (8)의 $Count_{dip}(gram_n)$ 은 예측 문장에서 특정 n-gram이 생성된 개수가 실제 정답 문장에서 등장한 개수보다 많을 경우, $Count_{match}(gram_n)$ 를 정답 문장에서 등장한 최대값으로 클리핑(clipping)하여 중복되는 단어에 대한 점수의 증가를 방지한다. 또한 식 (9)와 같이 예측 문장이 정답 문장보다 짧을 경우 1보다 작은 값을, 정답 문장보다 길 경우 1의 가중치를 부여한다. 최종적인 BLEU는 식 (10)과 같이 uni-gram부터 n-gram까지의 조화 평균으로 계산된다. w_n 는 식 (8)에 대한 가중치를 나타내며 일반적으로 $\frac{1}{n}$ 로

설정하여 모든 p_i 에 동일한 가중치를 적용한다.

$$p_i = \frac{\sum_{S \in \{References\}} \sum_{gram_n \in S} Count_{dip}(gram_n)}{\sum_{S \in \{Candidates\}} \sum_{gram_n \in S} Count(gram_n)} \quad (8)$$

$$BP = \begin{cases} 1 & \text{if candidate length} > \text{reference length} \\ e^{-\frac{\text{referencelength}}{\text{candidate length}}} & \text{if candidate length} \leq \text{reference length} \end{cases} \quad (9)$$

$$BLEU_N = BP \cdot \exp(\sum_{n=1}^N w_n \log p_n) \quad (10)$$

4.2 실험 결과

본 연구는 제안 방법론의 효과를 입증하기 위해 3개의 국방 M&S로부터 나온 8개의 전장 상황 인지 보고서 데이터에 (reconnaissance wing video interpretation, engagement result, ground forces observation, legion UAV video interpretation, anti-aircraft tracking, radar detection, engagement situation, warship observation) 대해 실험을 진행하였고, 제안 방법론의 우수성을 입증하기 위해 다양한 모델을 TaT4와 비교하였다. Structure-aware Seq2Seq(Liu *et al.*, 2018)은 단일 정보 테이블을 위한 전처리 방법, 듀얼 어텐션 메커니즘, 그리고 필드 게이트 인코더를 제안한 대표적인 table-to-text 방법론으로 제안 방법론의 비교 모델로 선정하였다. 그리고 사전 학습과 멀티 태스크 학습의 효과를 확인하기 위해서 사전 학습과 멀티 태스크를 모두 사용하지 않은 모델인 트랜스포머와 사전 학습 후 개별 보고서에 대해서만 전이 학습을 수행한 TaT4 without (w/o) mul-

Table 2. Comparison Results for Eight Battlefield Situation Awareness Reports

M&S	CJ21		CJ21		CJ21	
Report	Reconnaissance Wing Video Interpretation		Engagement Result		Ground Forces Observation	
Metric	BLEU-4	ROUGE-4 F1	BLEU-4	ROUGE-4 F1	BLEU-4	ROUGE-4 F1
Structure-Aware Seq2Seq	7.33±0.85	6.39±0.7	0.0±0.0	0.65±0.65	1.21±1.21	5.92±3.18
Transformer	8.66±2.05	4.9±0.73	17.69±4.88	6.75±1.0	31.27±4.95	24.98±3.97
TaT4 w/o MT	65.65±1.24	70.83±1.32	0.0±0.0	0.08±0.08	1.14±1.14	0.73±0.73
TaT4 (Proposed)	66.92±0.73	71.68±1.41	62.19±4.94	64.44±10.29	72.37±2.33	75.08±1.55
M&S	CJ21		CG		CG	
Report Type	Legion UAV Video Interpretation		Anti-aircraft Tracking		Radar Detection	
Metric	BLEU-4	ROUGE-4 F1	BLEU-4	ROUGE-4 F1	BLEU-4	ROUGE-4 F1
Structure-Aware Seq2Seq	12.93±1.94	20.83±4.16	0.0±0.0	0.0±0.0	0.06±0.06	1.82±0.58
Transformer	47.62±1.03	40.63±1.49	20.03±4.97	7.26±2.91	27.73±6.58	7.7±3.6
TaT4 w/o MT	26.44±13.24	17.47±8.75	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0
TaT4 (Ours)	66.95±0.97	69.72±2.18	39.29±7.13	35.62±6.45	66.27±0.94	59.92±2.12
M&S	CH		CH		CH	
Report Type	Engagement Situation		Warship Observation		CH	
Metric	BLEU-4	ROUGE-4 F1	BLEU-4	ROUGE-4 F1	BLEU-4	ROUGE-4 F1
Structure-Aware Seq2Seq	0.0±0.0	0.0±0.0	22.06±13.69	14.69±10.8		
Transformer	72.0±4.7	61.86±9.77	35.12±6.55	21.45±5.03		
TaT4 w/o MT	0.0±0.0	0.0±0.0	85.58±0.14	98.4±0.25		
TaT4 (Ours)	88.11±1.48	100.0±0.0	86.44±0.27	99.17±0.15		

ti-task (MT)를 제안하는 TaT4와 비교하였다. 비교 방법론 모두 멀티 태스크를 사용하지 않기 때문에 보고서별로 총 8개의 모델을 구축하여 성능을 평가하였다. <Table 2>는 제안 방법론과 비교 방법론들의 3회 반복 실험에 대한 평균 ROUGE-4 F1, BLEU-4와 표준오차를 나타낸 것이고, <Table 3>은 학습 데이터가 가장 많은 보고서와 가장 적은 보고서의 각 방법론들로부터 생성된 보고서 예시이다.

<Table 2>의 실험 결과에서 알 수 있듯이 모든 보고서에서 TaT4가 비교 방법론에 비해 매우 좋은 성능을 보이는 것을 확인할 수 있었다. 특히 structure-aware Seq2Seq은 대부분의 보고서에서 매우 낮은 성능을 보였으며 학습에 실패했다는 것을 알 수 있었다. Structure-aware Seq2Seq은 long short-term memory (LSTM) 기반 순환신경망을 기본 구조로 사용하기 때문에 전장 상황 보고서와 같이 보고서의 길이가 길어질수록 장기 의존성(long-term dependency)에 의해 학습이 제대로 되지 않는다. 실제로 <Table 3>에서 structure-aware Seq2Seq으로 생성된 보고서를 보면 트랜스포머 기반 모델들에 비해 생성된 시퀀스가 짧은 것을 알 수 있다. 특히 데이터가 59개 밖에 되지 않는 대공 항적 보고서 같은 경우, 학습에 완전히 실패한 것을 확인할 수 있다. 반면에 제안 모델인 TaT4는 셀프 어텐션 메커니즘 기반 트랜스포머를 기본 구조로 활용하고 동시에 사전 학습을 활용함으로써 장기 의존성에 문제점을 줄이고 안정된 학습을 할 수 있었다. 또한 멀티 태스크를 통해 모든 보고서를 학습 데이터로 사용하여 데이터 부족 문제를 극복할 수 있었다. 실제로 TaT4는 모든 보고서에서 적절한 문장을 생성하였

으며 데이터가 적은 상황에서도 의미 있는 수준의 보고서를 생성하는 것을 확인하였다. 특히 TaT4는 다중 정보 테이블을 입력값으로 가지는 대공 항적 보고서에서 다른 비교 방법론들보다 높은 BLEU-4와 ROUGE-4 F1 값을 보여주었다. 이를 통해 제안하는 전처리 기법과 멀티 태스크 학습이 다중 정보 테이블에서도 효과적이라는 것을 입증하였다.

사전 학습과 멀티 태스크가 TaT4 학습에 어떠한 영향을 주었는지 파악하기 위해서 사전 학습과 멀티 태스크를 모두 사용하지 않은 트랜스포머와 멀티 태스크를 사용하지 않은 TaT4 w/o MT 모델을 제안하는 TaT4와 비교하였다. <Table 2>에서 트랜스포머와 TaT4 w/o MT는 대부분 structure-aware Seq2Seq와 비슷하거나 더 좋은 성능을 보였다. 하지만 제안 방법론인 TaT4에서 멀티 태스크 학습을 제외하고 사전 학습만을 활용한 TaT4 w/o MT가 대부분 보고서에서 학습에 실패하였고 사전 학습을 사용하지 않은 트랜스포머보다도 성능이 낮음을 확인할 수 있었다. 반면에 전투 함정 관측 보고서(warship observation)와 정찰 비행단 영상 관측 보고서(reconnaissance wing video interpretation)에서는 BLEU-4 기준 65.65%, 85.58%를 달성하여 제안하는 TaT4에 준하는 성능을 보여주었다. 주목할 점은 학습에 실패한 보고서들의 데이터 개수이다. TaT4 w/o MT가 성공적으로 학습한 보고서는 CH warship observation과 CJ21 reconnaissance wing video이며 데이터 개수는 각각 약 7,000개와 3,400개로 학습하기 적당한 개수라고 판단할 수 있다. 하지만 실패한 보고서의 데이터 수는 500개 이하로 매우 적었다. 즉 학습에 적당한 데이터가 있을 때는 성능이

Table 3. Examples of Label and Generated Reports for Large Datasets

Report	CH Warship Observation	CG Anti-Aircraft Tracking
# of Data	6995	59
Label	해군기지 SHR058 상황보고합니다. 2022년 05월 26일 06시 41분 현재 제3국 수상함 제3국선박01이 고성 지역(38.974372N, 128.047309E)에서 남동쪽방향으로 15KTS로 이동하고 있습니다.	2022년 04월 05일 13시 11분 ~ 2022년 04월 05일 13시 12분 간 탐지된 항적은 총 전투기 MIG 23 2대였음. 전투기 MIG 23 2대는 판교지역(38.590556N, 126.898333E) - 판교지역(38.500000N, 126.833333E)를 경유하며 평균 420knot로 비행함
Structure-Aware Seq2Seq	해군기지 SHR058 상황보고. 2022년 05월 26일 06시 41분 현재 제3국 수상함 제3국선박01이 통천 지역(38.	2022일.월월
Transformer	해군기지 SHR058 상황보고합니다. 2022년 05월 26일 07시 17분 현재 제3국 수상함 제3국선박01이 고성 지역(38.834510N, 128.17E)에서 남동쪽방향으로 15KTS로 이동하고 있습니다.	탐지된 항적은 총 전투기 MIG 21DF 6대입니다. 2022년 04월 05일 08시 04월 05일 07시 04월 05일 08시 04월 05일 08시 04월 05일 08시 04월 05일
TaT4 w/o MT	해군기지 SHR058 상황보고. 2022년 05월 26일 06시 41분 현재 제3국 수상함 제3국선박01이 고성 지역(38.974372N, 128.047309E)에서 남동쪽방향으로 15KTS로 이동중	아에原原原原原原原原 때문이었다原原原原천~2 왜냐하면 김신회 왜냐면운동 여러분이 사람에게는戰(100 중이던 말이오세상을 것이오
TaT4	해군기지 SHR058 상황보고합니다. 2022년 05월 26일 06시 41분 현재 제3국 수상함 제3국선박01이 고성 지역(38.974372N, 128.047309E)에서 남동쪽방향으로 15KTS로 이동하고 있습니다.	탐지된 항적은 총 전투기 MIG 23 2대. 2022년 04월 05일 13시 12분부터 2022년 04월 05일 13시 12분까지 전투기 MIG 23 2대는 판교지역(38.590556N, 126.898333E)으로 평균 420knot로 비행함

향상되지만 그렇지 않을 경우 오히려 악영향을 끼치고 있다는 것을 알 수 있었다. 이는 신문 혹은 메신저 데이터로 이루어진 사전 학습 데이터와 전장 상황 인지 보고서 데이터의 도메인 차이로 인해 일반화가 제대로 되지 못했다고 해석할 수 있다. 실제로 <Table 3>에서 TaT4 w/o MT로 생성한 대공 항적 보고서는 보고서에 적합하지 않은 단어를 다수 출력한 것을 확인할 수 있었다. 하지만 TaT4 w/o MT와 달리 제안 방법론인 TaT4는 보고서 유형을 추가적으로 모델에 입력함으로써 여러 보고서를 단일 모델로 생성할 수 있는 멀티 태스크 학습을 수행하였다. 이는 모든 보고서를 하나의 데이터셋으로 취급하기 때문에 특정 보고서에 과적합 되지 않는 효과와 데이터 증가 효과를 통해 일반화 성능을 향상시킬 수 있었다. 결론적으로 실험 결과들을 통해 사전 학습과 멀티 태스크의 상호 보완적인 관계가 TaT4의 성능을 향상시킬 수 있었고 TaT4가 기존 table-to-text 모델보다 우수하다는 것을 입증하였다.

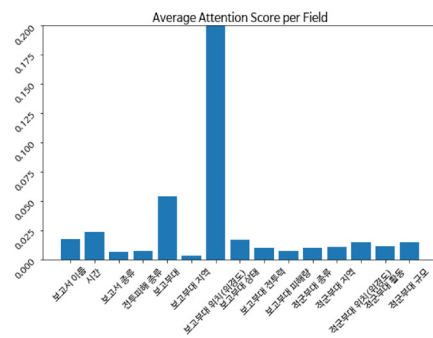
또한 실제 전장 상황에서는 모든 필드 정보를 얻을 수 없기 때문에 하나의 보고서에서도 수집된 정보와 현 상황에 맞게 적절한 보고서를 생성할 수 있어야 한다. <Figure 8>은 교전 결과 보고서에서 적군 정보 유무에 따른 데이터 예시(table)와 예측값(prediction)을 나타낸다. <Figure 8>의 어텐션 스코어는 TaT4가 적군 정보의 유무에 따라 어떤 필드에 가중치를 두었는지 파악하기 위해 각 필드에 해당하는 토큰들의 어텐션 스코어를 산출한 값이다. 전장 상황 인지 보고서에서 가장 중요한 부분은 상황이 발생한 정확한 시간과 장소를 보고하는 것이다. 본 연구는 <Figure 8>을 통해 두 보고서 모두 보고 부대

의 위치(위경도)에 가중치가 높은 것을 알 수 있었다. 특히 두 번째 그림의 경우에는 적군 부대를 탐지한 상황으로 적군 부대의 규모와 종류, 이동 방향에 가중치가 높은 것을 알 수 있었다. 어텐션 스코어 결과는 TaT4가 상황에 맞게 적절한 보고서를 생성할 수 있음을 보였으며 실제 전장 상황에서 활용할 수 있는 가능성을 보여주었다.

5. 결론

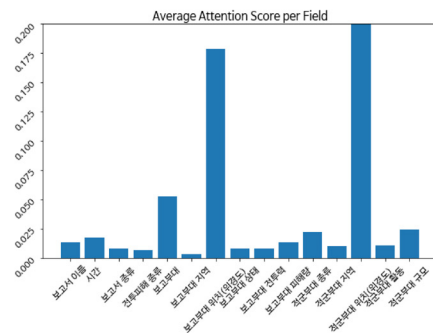
본 연구는 지능형 전장 상황 인식을 위한 서비스 구축에 앞서 전장 상황 로그 데이터로부터 정확한 정보를 담고 있는 비정형 전장 상황 인지 보고서 생성을 위한 text-to-text 멀티 태스크 학습 프레임워크를 제안하였다. 전장 상황 로그 데이터를 확보하기 위해 3개의 국방 M&S에서 시나리오를 입력하고 시뮬레이션을 진행하여 초기 로그를 수집 및 보고서별로 데이터셋을 구축하였다. 제안 방법론인 TaT4는 전장 상황 로그 데이터의 구조적 특징을 반영하기 위해 스페셜 토큰을 활용한 table-to-text 전처리 기법과 도메인이 서로 다른 데이터로 사전 학습한 모델 기반으로 멀티 태스크 학습을 수행하여 일반화 성능을 향상시켰다. 특히 제안 전처리 기법과 멀티 태스크 학습이 단일 정보 테이블 뿐만 아니라 다중 정보 테이블에도 적용 가능함을 실험적으로 보여주었다. 또한 TaT4가 적군과 아군의 상태, 정보의 유무가 다양한 전장 상황에서 어떠한 필드가 중요한 정보인지를 제공하였다. 본 연구는 비용적/공간적

시간	2019년 04월 05일 04시 04분
보고서 종류	적 접촉 단절
전투피해 종류	X
보고부대	A사단GOP1대대
보고부대 지역	개성 지역
보고부대 위경도	37.878508N, 126.713415E
보고부대 피해량	X
...	...
적군부대 종류	X
적군부대 지역	X
적군부대 위경도	X
적군부대 활동	X
적군부대 규모	X



2019년 04월 05일 04시 04분 개성 지역(37.878508N, 126.713415E)에 위치한 아군 A사단GOP1대대가 적과 접촉을 단절함. 현재 부대는 철수중이며, 전투력은 45%입니다.

시간	2019년 04월 05일 08시 30분
보고서 종류	적 접촉 상황
전투피해 종류	X
보고부대	A사단2GP
보고부대 지역	개성 지역
보고부대 위경도	37.938559N, 126.696763E
보고부대 피해량	X
...	X
적군부대 종류	경보병
적군부대 지역	개성 지역
적군부대 위경도	37.942027N, 126.689828E
적군부대 활동	남동 이동
적군부대 규모	250



2019년 04월 05일 08시 30분 개성 지역(37.942059N, 126.696726E)에 위치한 아군 A사단2GP는 적군 경보병부대의 접촉하였습니다. 관측된 적은 개성 지역(37.943027E)에서 남동쪽으로 이동하고 있습니다.

Figure 8. Comparison of Attention Scores Depending on Presence of Enemy Information on CJ21 Engagement Result

한계로 인해 실제 전장 환경을 모의하지 못하는 상황에서 국방 M&S를 활용하여 데이터를 생성하고 비정형 보고서를 생성하는 최초의 연구라는 점에서 의의가 있다. 또한 제안 방법론이 지식 베이스 등 지능형 전장 상황 인식 서비스 구축을 위한 비정형 보고서를 제공할 수 있을 것으로 기대한다.

본 연구에서는 멀티 태스크 학습 단계에서 각 보고서별 데이터셋의 크기를 고려하지 않고 학습 데이터로 활용하였다. 하지만 멀티 태스크 학습 시 개별 데이터셋에 대한 불균형이 학습 단계에서 어떠한 영향을 미치는지 파악할 필요가 있다. 또한 전장 상황 도메인 데이터에 대해 사전 학습 모델을 사용할 경우, 일반화 성능을 향상시키기 위한 적정 데이터 크기에 대한 고찰이 필요하다. 향후 연구에서는 멀티 태스크 학습에서 데이터 불균형이 학습에 어떠한 악영향을 미치는지를 파악하고 이를 해결하기 위해 오버샘플링(over-sampling)과 데이터셋 크기를 고려한 가중 손실 함수를 적용하고자 한다. 또한 데이터 개수에 따른 전이 학습 성능의 비교를 통해 정확한 정보를 내포하는 보고서를 생성하기 위한 적정 데이터 개수를 탐색하고자 한다. 제안 방법론은 생성된 보고서 내용에서 보고 및 타겟 부대의 위경도와 같이 고유 명사 혹은 숫자에서 일부 부정확한 정보를 산출하였다. 이를 보완하기 위해 향후 연구에서는 copy mechanism을 적용하여 입력 시퀀스로부터 숫자 정보 혹은 말뭉치에 없는 고유 명사를 정보 왜곡 없이 복사하여 출력 시퀀스를 생성하는 연구를 진행할 예정이다.

참고문헌

- Bao, J., Tang, D., Duan, N., Yan, Z., Lv, Y., Zhou, M., and Zhao, T. (2018), Table-to-text: Describing table region with natural language, In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 32, No. 1).
- Devlin, J., Chang, M. W., Lee, K., and Toutanova, K. (2018), Bert: Pre-training of deep bidirectional transformers for language understanding, arXiv preprint arXiv:1810.04805.
- Gu, J., Lu, Z., Li, H., and Li, V. O. (2016), Incorporating copying mechanism in sequence-to-sequence learning, arXiv preprint arXiv:1603.06393.
- Han, C. H., Cho, J. H., and Lee, S. K. (2011), A study for autonomous intelligence of computer-generated forces, *Journal of the Korea Society for Simulation*, **20**(1), 69-77.
- Han, C. and Lee, J. K. (2019), A methodology for defense AI command&control platform construction, *The Journal of Korean Institute of Communications and Information Sciences*, **44**(4), 774-781.
- Jin, S. Y. and Lee, W. S. (2020), A study on multiple reasoning technology for intelligent battlefield situational awareness, *The Journal of Korean Institute of Communications and Information Sciences*, **45**(6), 2020-6.
- Jo, S. H., Kim, H. J., Jin, S. Y., and Lee, W. S. (2020), A study on building knowledge base for intelligent battlefield awareness service, *Journal of the Korea Society of Computer and Information*, **25**(4), 11-17.
- Kim, Y. H., Song, Y. S., and Kim, C. O. (2019), A study on the interoperability of ROK air force virtual and constructive simulation, *Journal of the Korea Society for Simulation*, **28**(2), 169-177.
- Kim, S. M., Jin, S. Y., and Lee, W. S. (2021), A study on the extraction of similar information using knowledge base embedding for battlefield awareness, *Journal of the Korea Society of Computer and Information*, **26**(11), 33-40.
- Lin, C. Y. (2004), Rouge: A package for automatic evaluation of summaries, In *Proceedings of the 42th annual meeting of the Association for Computational Linguistics*, 74-81.
- Lee, K. H., Lim, H. J., and Kim, K. E. (2012), A case study on modeling computer generated forces based on factored POMDPs, In *Proceedings of the Korean Information Science Society Conference*, 333-335.
- Lebret, R., Grangier, D., and Auli, M. (2016), Neural text generation from structured data with application to the biography domain, arXiv preprint arXiv:1603.07771.
- Liu, T., Wang, K., Sha, L., Chang, B., and Sui, Z. (2018), Table-to-text generation by structure-aware seq2seq learning, In *Thirty-Second AAAI Conference on Artificial Intelligence*.
- Liu, X., He, P., Chen, W., and Gao, J. (2019), Multi-task deep neural networks for natural language understanding, arXiv preprint arXiv:1901.11504.
- Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., and Riedmiller, M. (2013), Playing atari with deep reinforcement learning, arXiv preprint arXiv:1312.5602.
- Mnih, V., Badia, A. P., Mirza, M., Graves, A., Lillicrap, T., Harley, T., ... and Kavukcuoglu, K. (2016), Asynchronous methods for deep reinforcement learning, In *International Conference on Machine Learning*, 1928-1937.
- Moon, H. S. (2018), A study on an expert system for close combat battlefield situation assessment in war-game models using machine learning, *Korean Journal of Military Arts and Science*, **74**(3), 315-335.
- Nallapati, R., Zhou, B., Gulcehre, C., and Xiang, B. (2016), Abstractive text summarization using sequence-to-sequence rnns and beyond, arXiv preprint arXiv:1602.06023.
- Papineni, K., Roukos, S., Ward, T., and Zhu, W. J. (2002), Bleu: A method for automatic evaluation of machine translation, In *Proceedings of the 40th annual meeting of the Association for Computational Linguistics*, 311-318.
- Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., ... and Liu, P. J. (2020), Exploring the limits of transfer learning with a unified text-to-text transformer, *Journal of Machine Learning Research*, **21**(140), 1-67.
- Shazeer, N. and Stern, M. (2018), Adafactor: Adaptive learning rates with sublinear memory cost, In *International Conference on Machine Learning*, 4596-4604.
- Shin, S., Kim, S., and Seo, H. T. (2021), Construction of bilingually pre-trained language model from large-scaled Korean and English corpus, In *Annual Conference on Human and Language Technology*, 419-422.
- Toghiani-Rizi, B., Kamrani, F., Luotsinen, L. J., and Gisslén, L. (2017), Evaluating deep reinforcement learning for computer generated forces in ground combat simulation, In *2017 IEEE International Conference on Systems, Man, and Cybernetics*, 3433-3438.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... and Polosukhin, I. (2017), Attention is all you need, *Advances in Neural Information Processing Systems*, 30.

저자소개

허종국 : 고려대학교 산업경영공학부에서 2021년 학사 학위를 취득하고, 고려대학교 산업경영공학과에서 석박통합과정에 재학 중이다. 연구분야는 Self-Supervised Learning, Graph Neural Networks, Reinforcement Learning이다.

임새린 : 고려대학교 산업경영공학부에서 2021년 학사 학위를 취득하고, 고려대학교 산업경영공학과에서 석박통합과정에 재학 중이다. 연구분야는 Time-Series Deep Learning, Self-Supervised Learning이다.

안시후 : 선문대학교 글로벌 소프트웨어학과에서 2021년 학사 학위를 취득하고, 고려대학교 산업경영공학과에서 석박통합과정에 재학 중이다. 연구분야는 Computer Vision, Pose estimation, Deep Learning for signal analysis이다.

박진혁 : 고려대학교 정보통계학과와 산업경영공학부에서 2019년 학사학위를 취득하고, 고려대학교 산업경영공학과에서 석박통합과정에 재학 중이다. 연구분야는 Unsupervised Semantic Segmentation, Object Detection이다.

이영재 : 고려대학교 정보통계학과와 산업경영공학부에서 2019년 학사학위를 취득하고, 고려대학교 산업경영공학과에서 박사과정에 재학 중이다. 연구분야는 Reinforcement Learning, Self-Supervised Learning이다.

조용원 : 고려대학교 정보통계학과와 산업경영공학부에서 2019년 학사학위를 취득하고, 고려대학교 산업경영공학과에서 박사과정에 재학 중이다. 연구분야는 Image segmentation algorithms and its applications, Anomaly detection and segmentation이다.

목충협 : 고려대학교 산업경영공학부에서 2019년 학사 학위를 취득하고, 고려대학교 산업경영공학과에서 박사과정에 재학 중이다. 연구분야는 Anomaly Detection, Self-Supervised Learning, Multi-task Learning이다.

조준호 : 동아대학교 컴퓨터공학과에서 2002년 학사학위를 취득하고, 2015년 한성대학교 국방M&S학으로 석사학위를 취득하였다. 2001년부터 리얼타임비주얼(주) 기술연구소에서 수석 연구원으로 재직 중이다. 주요 연구분야는 분산 시뮬레이션, Visualization이다.

김성범 : 고려대학교 산업경영공학부 교수로 2009년부터 재직하고 있으며, 인공지능공학연구소 소장 및 기업산학협력센터 센터장을 역임했다. 미국 University of Texas at Arlington 산업공학과에서 교수를 역임하였으며, 한양대학교 산업공학과에서 학사학위를 미국 Georgia Institute of Technology에서 산업공학 석사 및 박사학위를 취득하였다. 인공지능, 머신러닝, 최적화 방법론을 개발하고 이를 다양한 공학, 자연과학, 사회과학 분야에 응용하는 연구를 수행하고 있다.