



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0142818
(43) 공개일자 2024년10월02일

(51) 국제특허분류(Int. Cl.)
G01W 1/10 (2006.01) G01W 1/18 (2006.01)
G06N 3/04 (2023.01) G06N 3/092 (2023.01)
(52) CPC특허분류
G01W 1/10 (2013.01)
G01W 1/18 (2013.01)
(21) 출원번호 10-2023-0037331
(22) 출원일자 2023년03월22일
심사청구일자 2023년03월22일

(71) 출원인
경희대학교 산학협력단
경기도 용인시 기흥구 덕영대로 1732 (서천동, 경희대학교 국제캠퍼스내)
(72) 발명자
문용재
경기도 용인시 수지구 신봉2로 72, 209동 1404호 (신봉동, 신봉마을자이2차아파트)
이강우
경기 수원시 영통구 봉영로1482번길 7-4, 303호 (영통동)
(74) 대리인
유미특허법인

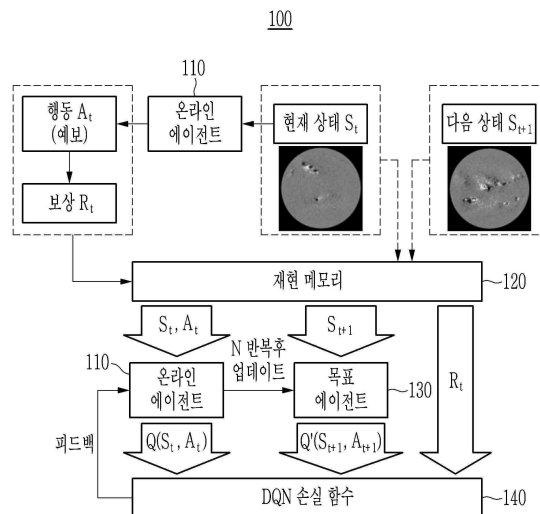
전체 청구항 수 : 총 14 항

(54) 발명의 명칭 딥러닝 강화학습을 이용한 태양 중요 플레어 예측 방법 및 그 장치

(57) 요약

딥러닝 강화학습을 이용한 태양 중요 플레어 예측 방법 및 그 장치가 개시된다. 이 방법은 컴퓨터로 구현되는 태양 중요 플레어 예측 장치에서 수행되는 태양 중요 플레어 예측 방법으로서, 복수의 태양 전면 자기장 지도에 대한 학습 데이터 또는 예측 대상 데이터를 수신하는 단계를 포함한다. 이 방법은 학습 데이터에 대해 심층 강화학습을 수행하여 태양 중요 플레어 예측 모델을 생성하는 단계를 더 포함한다. 이 방법은 태양 중요 플레어 예측 모델을 사용하여 예측 대상 데이터에 대한 태양 중요 플레어 발생 여부를 예측하는 단계를 더 포함한다. 이 방법에서, 심층 강화학습시 태양 중요 플레어에 대한 예측의 성공과 실패에 대한 보상이 서로 다르게 적용된다.

대표도 - 도1



(52) CPC특허분류

G06N 3/04 (2023.01)

G06N 3/092 (2023.01)

G01W 2201/00 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 1345341316

과제번호 2021R1A6A3A01088835

부처명 교육부

과제관리(전문)기관명 한국연구재단

연구사업명 박사후국내연수

연구과제명 설명 가능한 딥러닝과 비지도 딥러닝을 활용한 태양 자기장 지도의 새로운 과학적
분석 연구

기 여 율 1/1

과제수행기관명 경희대학교 산학협력단

연구기간 2021.09.01 ~ 2024.08.31

명세서

청구범위

청구항 1

컴퓨터로 구현되는 태양 중요 플레이어 예측 장치에서 수행되는 태양 중요 플레이어 예측 방법으로서,
복수의 태양 전면 자기장 지도에 대한 학습 데이터 또는 예측 대상 데이터를 수신하는 단계,
상기 학습 데이터에 대해 심층 강화학습을 수행하여 태양 중요 플레이어 예측 모델을 생성하는 단계, 그리고
상기 태양 중요 플레이어 예측 모델을 사용하여 상기 예측 대상 데이터에 대한 태양 중요 플레이어 발생 여부를 예측하는 단계
를 포함하며,
상기 심층 강화학습시 태양 중요 플레이어에 대한 예측의 성공과 실패에 대한 보상을 서로 다르게 적용하는,
태양 중요 플레이어 예측 방법.

청구항 2

제1항에 있어서,
상기 심층 강화학습은 DQN(Deep Q-Network) 알고리즘 또는 더블(double) DQN 알고리즘인,
태양 중요 플레이어 예측 방법.

청구항 3

제2항에 있어서,
상기 보상은 상기 태양 중요 플레이어에 대한 긍정 예측(Positive) 및 부정 예측(Negative)과 상기 태양 중요 플레이어 예측 모델의 예측 결과인 긍정 결과 및 부정 결과에 따라 4가지 보상(TP(True Positive), FP(False Positive), TN(True Negative), FN(False Negative))으로 주어지며,
상기 4가지 보상은 서로 다른 보상값을 갖도록 설정되는,
태양 중요 플레이어 예측 방법.

청구항 4

제3항에 있어서,
상기 4가지 보상 중 한가지 보상의 값을 1로 설정하고, 나머지 보상에 대한 값은 1에 대한 서로 다른 비율로 설정되는,
태양 중요 플레이어 예측 방법.

청구항 5

제2항에 있어서,
상기 학습 데이터에 대해 심층 강화학습을 수행하여 태양 중요 플레이어 예측 모델을 생성하는 단계는,
현재 분석하고자 하는 상태인 현재 상태, 상기 현재 상태를 온라인 에이전트에 입력하여 분석했을 때의 결과인 행동, 상기 행동으로 인한 보상, 및 무작위로 선정된 다음 상태를 포함하는 4개의 데이터를 한 세트로서 복수의 데이터 세트를 재현 메모리에 저장하는 단계,
상기 재현 메모리에 저장된 데이터 세트를 이용하여 상기 온라인 에이전트를 훈련시키고, 훈련된 온라인 에이전트의 학습 정보를 미리 설정된 업데이트 시간에 따라 정기적으로 목표 에이전트에게 복사하는 단계, 그리고

상기 목표 에이전트를 상기 태양 중요 플레이어 예측 모델로 설정하는 단계를 포함하는, 태양 중요 플레이어 예측 방법.

청구항 6

제5항에 있어서,
상기 업데이트 시간은 1 에포크(epoch) 또는 2 에포크인,
태양 중요 플레이어 예측 방법.

청구항 7

제5항에 있어서,
상기 온라인 에이전트 및 상기 목표 에이전트는 각각 CNN(Convolutional Neural Network) 알고리즘으로 구현되는,
태양 중요 플레이어 예측 방법.

청구항 8

제6항에 있어서,
보상 유형의 개수, DQN 알고리즘 또는 더블 DQN 알고리즘, 및 2개의 업데이트 시간에 따라 생성되는 복수의 태양 중요 플레이어 예측 모델에 대해 HSS(Heidke Skill Score), F1, TSS(True Skill Statistic) 및 ApSS(Appleman's Skill Score)의 기술 점수에 따른 평가에 기초하여 최적의 태양 중요 플레이어 예측 모델이 결정되는,
태양 중요 플레이어 예측 방법.

청구항 9

태양 중요 플레이어 예측 장치로서,
복수의 태양 전면 자기장 지도에 대한 학습 데이터 또는 예측 대상 데이터를 수신하는 데이터 수신부,
상기 학습 데이터에 대해 심층 강화학습을 수행하여 태양 중요 플레이어 예측 모델을 생성하는 학습부, 그리고
상기 태양 중요 플레이어 예측 모델을 사용하여 상기 예측 대상 데이터에 대한 태양 중요 플레이어 발생 여부를 예측하는 예측부를 포함하며,
상기 학습부는 상기 심층 강화학습시 태양 중요 플레이어에 대한 예측의 성공과 실패에 대한 보상을 서로 다르게 적용하는,
태양 중요 플레이어 예측 장치.

청구항 10

제9항에 있어서,
상기 학습부는,
재현 메모리, 온라인 에이전트 및 목표 에이전트를 포함하는 DQN 또는 더블 DQN 알고리즘을 사용하여 상기 태양 중요 플레이어 예측 모델을 생성하고,
현재 분석하고자 하는 상태인 현재 상태, 상기 현재 상태를 상기 온라인 에이전트에 입력하여 분석했을 때의 결과인 행동, 상기 행동으로 인한 보상, 및 무작위로 선정된 다음 상태를 포함하는 4개의 데이터를 한 세트로서 복수의 데이터 세트를 상기 재현 메모리에 저장하며,
상기 재현 메모리에 저장된 데이터 세트를 이용하여 상기 온라인 에이전트를 훈련시키고,

훈련된 온라인 에이전트의 학습 정보를 미리 설정된 업데이트 시간에 따라 정기적으로 상기 목표 에이전트에게 복사하여 상기 태양 중요 플레어 예측 모델을 생성하는,

태양 중요 플레어 예측 장치.

청구항 11

제10항에 있어서,

상기 업데이트 시간은 1 에포크 또는 2 에포크인,

태양 중요 플레어 예측 장치.

청구항 12

제10항에 있어서,

상기 보상은 상기 태양 중요 플레어에 대한 긍정 예측 및 부정 예측과 상기 태양 중요 플레어 예측 모델의 예측 결과인 긍정 결과 및 부정 결과에 따라 4가지 보상(TP(True Positive), FP(False Positive), TN(True Negative), FN(False Negative))으로 주어지며,

상기 4가지 보상은 서로 다른 보상값을 갖도록 설정되는,

태양 중요 플레어 예측 장치.

청구항 13

제12항에 있어서,

상기 4가지 보상 중 한가지 보상의 값을 1로 설정하고, 나머지 보상에 대한 값은 1에 대한 서로 다른 비율로 설정되는,

태양 중요 플레어 예측 장치.

청구항 14

제10항에 있어서,

보상 유형의 개수, DQN 알고리즘 또는 더블 DQN 알고리즘, 및 2개의 업데이트 시간에 따라 생성되는 복수의 태양 중요 플레어 예측 모델에 대해 HSS, F1, TSS 및 ApSS의 기술 점수에 따른 평가에 기초하여 최적의 태양 중요 플레어 예측 모델이 결정되는,

태양 중요 플레어 예측 장치.

발명의 설명

기술 분야

[0001] 본 발명은 딥러닝 강화학습을 이용한 태양 중요 플레어 예측 방법 및 그 장치에 관한 것이다.

배경 기술

[0002] 태양 플레어(solar flare)는 태양에서 갑자기 번쩍이는 것으로, 광범위한 방출 스펙트럼에서 막대한 양의 에너지를 방출하고 입자를 행성 간 공간으로 가속시킨다. 이러한 플레어는 soft X-ray 플럭스(0.1-0.8nm)에서의 강도에 따라 5개의 클래스(A, B, C, M, X)로 분류된다. 각각의 클래스는 10배의 에너지 증가를 나타내며, 강한 X선 플럭스와 강한 플레어에서 가속된 입자는 위성 항력, GPS 중단, 라디오 페이드아웃과 같은 막대한 경제적 손실을 초래한다.

[0003] 태양 플레어 발생률은 에너지에 따라 다르다. 중요 플레어(또는 "강한 플레어")($\geq M$ -class)는 드물게 발생하고 약한 플레어($\leq C$ -class)는 자주 발생한다.

[0004] 기존의 중요 플레어 예측(또는 "예보")은 "불균형 데이터셋 문제(Imbalanced dataset problem)"로 인하여 성능 저하 문제를 겪고 있다. 여기서, 불균형 데이터셋 문제란, 예측 알고리즘이 드물게 발생하는 이벤트(즉, 중

요 플레어 발생 이벤트)보다 흔하게 발생하는 이벤트를 맞추는데 과적합되는 현상을 말한다.

- [0005] 대부분의 경우 드물게 발생하는 현상을 예측하는 것이 예측 모델의 목적이라는 것을 고려하면, 불균형 데이터셋 문제는 예측 모델의 성능 향상을 위하여 반드시 극복되어야 되는 문제이다.
- [0006] 태양 플레어 예측 모델 또한 드물게 발생하는 중요 플레어의 예측에 있어 불균형 데이터셋 문제로 인하여 어려움을 겪고 있다.
- [0007] 한편, 기존의 인공지능 강화학습 모델들은 모델 학습 과정에서 고정된 크기의 보상을 받도록 설계되어 있다. 예를 들어, 블록 깨기 게임을 수행하는 강화학습 모델의 경우, 게임은 더 높은 곳의 블록을 부술수록 더 많은 점수를 얻도록 설계되어 있지만, 모델은 점수를 얻을 시 항상 고정된 값, 예를 들어 +1의 보상을 받도록 설계되어 있다.
- [0008] 따라서, 기존의 플레어 예측 모델이 불균형 데이터셋 예측 문제를 해결하기 위해서는 기존의 방식과 다른 차별화 된 방식의 적용이 요구된다.

발명의 내용

해결하려는 과제

- [0009] 본 발명이 해결하고자 하는 과제는 태양 중요 플레어의 발생을 정확하게 예측할 수 있는 딥러닝 강화학습을 이용한 태양 중요 플레어 예측 방법 및 그 장치를 제공하는 것이다.

과제의 해결 수단

- [0010] 상기한 바와 같은 본 발명의 과제를 달성하고, 후술하는 본 발명의 특징적인 효과를 실현하기 위한, 본 발명의 특징적인 구성은 하기와 같다.
- [0011] 본 발명의 일 측면에 따르면, 태양 중요 플레어 예측 방법이 제공되며, 이 방법은,
- [0012] 컴퓨터로 구현되는 태양 중요 플레어 예측 장치에서 수행되는 태양 중요 플레어 예측 방법으로서, 복수의 태양 전면 자기장 지도에 대한 학습 데이터 또는 예측 대상 데이터를 수신하는 단계, 상기 학습 데이터에 대해 심층 강화학습을 수행하여 태양 중요 플레어 예측 모델을 생성하는 단계, 그리고 상기 태양 중요 플레어 예측 모델을 사용하여 상기 예측 대상 데이터에 대한 태양 중요 플레어 발생 여부를 예측하는 단계를 포함하며, 상기 심층 강화학습시 태양 중요 플레어에 대한 예측의 성공과 실패에 대한 보상을 서로 다르게 적용한다.
- [0013] 여기서, 상기 심층 강화학습은 DQN(Deep Q-Network) 알고리즘 또는 더블(double) DQN 알고리즘이다.
- [0014] 또한, 상기 보상은 상기 태양 중요 플레어에 대한 긍정 예측(Positive) 및 부정 예측(Negative)과 상기 태양 중요 플레어 예측 모델의 예측 결과인 긍정 결과 및 부정 결과에 따라 4가지 보상(TP(True Positive), FP(False Positive), TN(True Negative), FN(False Negative))으로 주어지며, 상기 4가지 보상은 서로 다른 보상값을 갖도록 설정된다.
- [0015] 또한, 상기 4가지 보상 중 한가지 보상의 값을 1로 설정하고, 나머지 보상에 대한 값은 1에 대한 서로 다른 비율로 설정된다.
- [0016] 또한, 상기 학습 데이터에 대해 심층 강화학습을 수행하여 태양 중요 플레어 예측 모델을 생성하는 단계는, 현재 분석하고자 하는 상태인 현재 상태, 상기 현재 상태를 온라인 에이전트에 입력하여 분석했을 때의 결과인 행동, 상기 행동으로 인한 보상, 및 무작위로 선정된 다음 상태를 포함하는 4개의 데이터를 한 세트로서 복수의 데이터 세트를 재현 메모리에 저장하는 단계, 상기 재현 메모리에 저장된 데이터 세트를 이용하여 상기 온라인 에이전트를 훈련시키고, 훈련된 온라인 에이전트의 학습 정보를 미리 설정된 업데이트 시간에 따라 정기적으로 목표 에이전트에게 복사하는 단계, 그리고 상기 목표 에이전트를 상기 태양 중요 플레어 예측 모델로 설정하는 단계를 포함한다.
- [0017] 또한, 상기 업데이트 시간은 1 에포크(epoch) 또는 2 에포크이다.
- [0018] 또한, 상기 온라인 에이전트 및 상기 목표 에이전트는 각각 CNN(Convolutional Neural Network) 알고리즘으로 구현된다.
- [0019] 또한, 보상 유형의 개수, DQN 알고리즘 또는 더블 DQN 알고리즘, 및 2개의 업데이트 시간에 따라 생성되는 복수

의 태양 중요 플레어 예측 모델에 대해 HSS(Heidke Skill Score), F1, TSS(True Skill Statistic) 및 ApSS(Appleman's Skill Score)의 기술 점수에 따른 평가에 기초하여 최적의 태양 중요 플레어 예측 모델이 결정된다.

[0020] 본 발명의 다른 측면에 따르면, 태양 중요 플레어 예측 장치가 제공되며, 이 장치는,

[0021] 복수의 태양 전면 자기장 지도에 대한 학습 데이터 또는 예측 대상 데이터를 수신하는 데이터 수신부, 상기 학습 데이터에 대해 심층 강화학습을 수행하여 태양 중요 플레어 예측 모델을 생성하는 학습부, 그리고 상기 태양 중요 플레어 예측 모델을 사용하여 상기 예측 대상 데이터에 대한 태양 중요 플레어 발생 여부를 예측하는 예측부를 포함하며, 상기 학습부는 상기 심층 강화학습시 태양 중요 플레어에 대한 예측의 성공과 실패에 대한 보상을 서로 다르게 적용한다.

[0022] 또한, 상기 학습부는, 재현 메모리, 온라인 에이전트 및 목표 에이전트를 포함하는 DQN 또는 더블 DQN 알고리즘을 사용하여 상기 태양 중요 플레어 예측 모델을 생성하고, 현재 분석하고자 하는 상태인 현재 상태, 상기 현재 상태를 상기 온라인 에이전트에 입력하여 분석했을 때의 결과인 행동, 상기 행동으로 인한 보상, 및 무작위로 선정된 다음 상태를 포함하는 4개의 데이터를 한 세트로서 복수의 데이터 세트를 상기 재현 메모리에 저장하며, 상기 재현 메모리에 저장된 데이터 세트를 이용하여 상기 온라인 에이전트를 훈련시키고, 훈련된 온라인 에이전트의 학습 정보를 미리 설정된 업데이트 시간에 따라 정기적으로 상기 목표 에이전트에게 복사하여 상기 태양 중요 플레어 예측 모델을 생성한다.

발명의 효과

[0023] 본 발명에 따르면, 딥러닝 강화학습을 이용하여 태양 중요 플레어의 발생을 정확하게 예측할 수 있다.

[0024] 또한, 태양 플레어 발생을 정확히 예측함으로써 인공위성 운영자, GPS 사용자, 우주비행사, 극항로 노선 비행사 등 관련 산업 종사자에게 도움이 될 수 있다.

[0025] 따라서, 경제적, 산업적 피해를 최소화할 수 있다.

[0026] 또한, 여러 모델을 설계하지 않아도 주어진 데이터와 모델 구조만 사용하여 예보 용도에 특화된 모델을 제작할 수 있게 됨으로서 더욱 다양하고 신속한 모델 개발과 그에 따른 예측 분야의 발전을 기대할 수 있다.

도면의 간단한 설명

[0027] 도 1은 본 발명의 실시예에 따른 딥러닝 강화학습을 이용한 태양 중요 플레어 예측 모델을 훈련시키기 위한 개략적인 구성도이다.

도 2는 도 1에 도시된 온라인 에이전트 및 목표 에이전트를 구현하는 CNN 모델 아키텍처의 개략도이다.

도 3은 본 발명의 실시예에 따른 딥러닝 강화학습에서 사용되는 태양 자기장 지도 입력의 개략도이다.

도 4는 본 발명의 실시예에 따라 테스트 데이터 세트를 사용하여 모델 3을 기반으로 10개 모델의 평균 보상을 추적하는 훈련 곡선을 도시한다.

도 5는 본 발명의 실시예에 따른 딥러닝 강화학습을 이용한 태양 중요 플레어 예측 장치의 구성 블록도이다.

도 6은 본 발명의 실시예에 따른 태양 중요 플레어 예측 방법의 개략적인 흐름도이다.

도 7은 본 발명의 다른 실시예에 따른 태양 중요 플레어 예측 장치의 구성 블록도이다.

발명을 실시하기 위한 구체적인 내용

[0028] 아래에서는 첨부한 도면을 참고로 하여 본 발명의 실시예에 대하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 상세히 설명한다. 그러나 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시예에 한정되지 않는다. 그리고 도면에서 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다

[0029] 명세서 전체에서, 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다. 또한, 명세서에 기재된 "...부", "...기", "모듈" 등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드

웨어나 소프트웨어 또는 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다.

- [0030] 본 발명에서 설명하는 장치들은 적어도 하나의 프로세서, 메모리 장치, 통신 장치 등을 포함하는 하드웨어로 구성되고, 지정된 장소에 하드웨어와 결합되어 실행되는 프로그램이 저장된다. 하드웨어는 본 발명의 방법을 실행할 수 있는 구성과 성능을 가진다. 프로그램은 도면들을 참고로 설명한 본 발명의 동작 방법을 구현한 명령어(instructions)를 포함하고, 프로세서와 메모리 장치 등의 하드웨어와 결합하여 본 발명을 실행한다.
- [0031] 이하, 본 발명의 실시예에 따른 딥러닝 강화학습을 이용한 태양 중요 플레어 예측 방법에 대해 설명한다.
- [0032] 도 1은 본 발명의 실시예에 따른 딥러닝 강화학습을 이용한 태양 중요 플레어 예측 모델을 훈련시키기 위한 개략적인 구성도이다. 도 1에는 본 발명의 실시예와 관련이 있는 구성요소들만이 도시되어 있다. 따라서, 본 발명이 속하는 기술 분야에서 통상의 기술자라면 도 1에 도시된 구성요소들 외에 다른 범용적인 구성요소들이 더 포함될 수 있음을 알 수 있다.
- [0033] 도 1에 도시된 태양 중요 플레어 예측 모델은 딥러닝 인공지능 알고리즘 중 하나인 DQN(Deep Q-Network)에 기반하는 모델로서, 여기서 DQN은 Google DeepMind에 의해 발표된 딥러닝과 강화학습을 결합한 심층강화학습에서 기본이 되는 알고리즘이다.
- [0034] 도 1을 참조하면, 본 발명의 실시예에서 사용되는 DQN(100)에서, 현재 분석하고자 하는 상태, 즉 현재 상태(S_t)를 온라인 에이전트(110)에 넣어 분석했을 때의 결과(즉, 예측 또는 예보)인 행동(A_t)으로 인한 보상(R_t), 그리고 무작위로 선정된 다음 데이터인 다음 상태(S_{t+1})로 구성되는 4개의 데이터가 한 세트로서 재현 메모리(Replay Memory)(120) 공간에 미리 저장된다.
- [0035] 이러한 저장을 계속 반복하면 재현 메모리(120)에 충분한 데이터 세트가 저장되고, 이렇게 충분한 데이터 세트가 저장된 재현 메모리(120)에서 무작위로 선정한 데이터 세트를 이용하여 온라인 에이전트(110)를 훈련시키고, 훈련된 온라인 에이전트(110)의 학습 정보를 정기적으로 목표 에이전트(130)에게 복사해 주는 방법의 훈련을 반복하여 목표 에이전트(130)를 학습시킨다.
- [0036] 한편, 재현 메모리(120)에 저장된 데이터 세트의 수는 일정 이상이 되면 오래된 데이터 세트부터 제거되도록 구현된다. 온라인 에이전트(110)는 효율적인 학습을 위하여 일정 확률로 무작위 결과를 택하도록 설정되어 있고, 목표 에이전트(130)는 무작위성 없이 가장 좋은 분석 결과를 택하도록 설정되어 있다. 최종적으로, 태양 중요 플레어 예측 모델(100)로서 사용되는 것은 목표 에이전트(130)이다.
- [0037] 보다 구체적으로 설명하면, 온라인 에이전트(110)의 예측 결과인 $Q(S_t, A_t)$ 를 최대화하는 방향으로 학습하기 위해, 목표 에이전트(130)와 DQN 손실 함수(140)를 활용한다. 온라인 에이전트(110)는 재현 메모리(120)로부터 현재 상태(S_t)에 관한 정보와 행동(A_t)에 관한 정보를 입력받는다. 따라서, 온라인 에이전트(110)는 현재 상태(S_t), 행동(A_t) 및 보상(R_t)에 관한 정보를 토대로 가치함수($Q(S_t, A_t)$)를 생성한다.
- [0038] 목표 에이전트(130)는 현재 상태(S_t)에서의 행동들 중 가장 가치함수가 최대화되는 방향의 행동을 수행하도록 한다. 목표 에이전트(130)는 재현 메모리(120)에서 얻은 다음 상태(S_{t+1}) 정보를 토대로 가장 가치가 높은 행동에 대한 가치함수($Q'(S_{t+1}, A'_{t+1})$)를 출력한다.
- [0039] 그리고, DQN 손실 함수(140)가 최소화되는 방향으로 온라인 에이전트(110)가 학습된다. 이 때, 학습되는 가중치는 목표값과의 차이, 즉 보상(A_t)과 가치함수($Q(S_t, A_t)$), 그리고 가치함수($Q(S_t, A_t)$)가 최대가 되는 가치함수($Q'(S_{t+1}, A'_{t+1})$)와의 평균제곱오차(mean square error, MSE)로서 업데이트될 수 있다.
- [0040] 여기서 가치함수란 상태(S_t)의 가치를 수치화한 함수로서, 상태(S_t)를 행동(A_t)으로 치환하기 위한 함수를 의미한다.
- [0041] 한편, 온라인 에이전트(110)는 ϵ -그리디(greedy) 정책을 사용하며 현재 상태(S_t)에서 행동(A_t)의 가치함수의 값(Q 값)을 계산한다. 반면에, 목표 에이전트(130)는 그리디 정책을 사용하여 다음 상태(S_{t+1})의 행동(A_{t+1})의 가치함수의 값(Q 값)을 계산한다. 여기서, 그리디 정책은 가장 가치 있는 행동을 선택하는 정책이고, ϵ -그리디 정책은 ϵ 의 확률로 랜덤 행동을 선택하고 $(1-\epsilon)$ 의 확률로 가장 가치 있는 행동을 선택하도록 하는 정책을 의미한다. 또한, 온라인 에이전트(110)는 모든 훈련 단계에서 업데이트되는 반면, 목표 에이전트(130)는 N 번 반복

후 온라인 에이전트(110)의 복제본으로 업데이트된다.

[0042] 전술한 DQN에서의 가치함수는 다음의 [수학식 1]과 같이 산출될 수 있다.

[0043] [수학식 1]

$$Q^O(S_t, A_t) = Q^O(S_t, A_t^O) + \alpha(R_{t+1} + \gamma Q^T(S_{t+1}, A_{t+1}^T) - Q^O(S_t, A_t^O))$$

[0045] 여기서, Q^O 는 온라인 에이전트(110)의 가치함수를 나타내고, Q^T 는 목표 에이전트(130)의 가치함수를 나타내며, α 는 학습률이고, γ 는 추정된 미래 가치에 대한 디스카운트 팩터이다.

[0046] 또한, 전술한 DQN 손실 함수(140)는 다음의 [수학식 2]와 같이 나타낼 수 있다.

[0047] [수학식 2]

$$L_{DQN} = (R_t + \gamma Q^T(S_{t+1}, A_{t+1}^T) - Q(S_t, A_t))^2$$

[0049] 여기서, γ 는 추정된 미래 가치에 대한 디스카운트 팩터이고, Q' 은 Q^T 와 동일하고, A' 은 A^T 와 동일하다.

[0050] 한편, 본 발명의 실시예에 따른 DQN에서는 온라인 에이전트(110) 및 목표 에이전트(130)에 대해 최적의 행동을 결정하기 위해 CNN(Convolutional Neural Network) 에이전트로서 구현한다. 여기서, CNN은 분류를 위한 2D 이미지 처리에 널리 사용되는 심층 신경망으로서, 입력 데이터의 값과 구조를 분석하기 위해 서로 다른 가중치를 갖고 서로 복잡하게 연결된 많은 컨볼루션 필터를 사용한다.

[0051] 본 발명의 실시예에서 사용되는 CNN 모델 아키텍처는 도 2에 도시되어 있다.

[0052] 도 2를 참조하면, 본 발명의 실시예에 따른 CNN 모델(200)은 초기 블록(210), 밀집 연결을 사용하는 5개의 밀집 블록(Dense block)(220) 및 최종 블록(230)으로 구성된다. 이 때, CNN 모델(200)에서, K는 커널 크기이고, D는 특징 차원의 개수이다.

[0053] 초기 블록(210)은 3×3 컨볼루션 커널(Conv)과 2×2 최대 풀링 레이어(Max pooling)로 구성된다.

[0054] 5개의 밀집 블록(220)은 동일한 구조를 가지며, 1개의 밀집 블록(220)에는 배치 정규화(Batch Normalization, BN)(212), 정류 선형 유닛(Rectified Linear Unit, ReLU)(222), 1×1 컨볼루션 커널(223), BN(224), ReLU(225), 3×3 컨볼루션 커널(226), 연결 레이어(227) 및 2×2 평균 풀링(Average Pooling) 레이어(228)가 순서대로 포함된다.

[0055] 최종 블록(230)은 BN(231), 2×2 평균 풀링 레이어(232) 및 완전 연결(Fully Connected) 레이어(233)로 구성된다.

[0056] 최종 블록의 두 출력은 각각 플레이어 및 비 플레이어에 대한 점수(241, 242)이다. 모델 예측의 결과로서 더 높은 점수가 사용된다.

[0057] 본 발명의 실시예에 따른 CNN 모델(200)은 DQN 방법의 적용 없이 그 자체로서도 태양 플레이어 예측 모델로서 작동하며, 이러한 경우 평균 제곱 오차 손실 함수를 사용하며, 그 손실 함수는 다음의 [수학식 3]과 같이 나타낼 수 있다.

[0058] [수학식 3]

$$L_{CNN} = (x_i - y_i)^2$$

[0060] 여기서, x_i 는 관측치를 나타내고, y_i 는 모델 결과를 나타낸다.

[0061] 전술한 CNN 모델(200)은, 예를 들어 도 3에 도시된 바와 같은 태양 자기장 지도가 입력으로 주어지면 CNN을 사용하여 태양 자기장 지도를 분석하여 태양 플레이어 및 비 플레이어에 대한 점수를 출력할 수 있다.

[0062] 전술한 CNN 모델(200)의 구조에 대한 세부 내용은 이미 잘 알려져 있으므로, 여기에서는 구체적인 설명을 생략한다.

[0063] 한편, 본 발명의 실시예에서는 전술한 DQN(200) 대신에 더블(Double) DQN을 적용할 수 있다. 여기서, 더블 DQN

은 더블 Q-러닝의 딥러닝 애플리케이션으로서, DQN의 문제점 중 하나인 Q-러닝이 다음 상태의 값을 결정하기 위해 최대값을 갖는 행동을 사용하기 때문에 행동 값을 과대 평가한다는 문제점을 피하기 위해 개발된 심층 강화 학습 알고리즘이다.

이러한 더블 DQN의 가치함수는 다음의 [수학식 4]와 같이 나타낼 수 있다.

[수학식 4]

$$Q^O(S_t, A_t) \leftarrow Q^O(S_t, A_t^O) + \alpha(R_t + \gamma Q^T(S_{t+1}, A_{t+1}^{O'}) - Q^O(S_t, A_t^O))$$

여기서, Q^O 는 온라인 에이전트(110)의 가치함수를 나타내고, Q^T 는 목표 에이전트(130)의 가치함수를 나타내고, A^O 와 $A_{t+1}^{O'}$ 은 각각 ϵ -그리디 정책과 그리디 정책으로 온라인 에이전트(110)에 의해 선택된 행동을 나타낸다.

더블 DQN의 목표 에이전트(130)는 온라인 에이전트(110)가 선택한 행동 $A_{t+1}^{O'}$ 을 사용하여 다음 상태의 값을 추정한다.

전술한 더블 DQN의 손실 함수는 다음의 [수학식 5]와 같이 나타낼 수 있다.

[수학식 5]

$$L_{DoubleDQN} = (R_t + \gamma Q^T(S_{t+1}, A_{t+1}^{O'}) - Q^O(S_t, A_t^O))^2$$

한편, 본 발명의 실시예에서는 전술한 DQN(100)(또는 더블 DQN)에서 TP(True Positive), FP(False Positive), TN(True Negative) 및 FN(False Negative) 각각에 대해 서로 다른 보상을 부여한다. 여기서, TP, FP, TN 및 FN은 다음의 [표 1]에 나타낸 바와 같이 정의될 수 있다.

표 1

		결과	
예측		Positive (긍정)	Negative (부정)
	Positive (긍정)	TP (True Positive)	FP (False Positive)
	Negative (부정)	FN (False Negative)	TN (True Negative)

따라서, 전술한 TP, FP, TN 및 FN에 대해, 예를 들어 다음의 [표 2]와 같이 서로 다른 보상이 부여될 수 있으며, 이러한 보상의 부여는 예시일 뿐이며 이들로만 제한되는 것은 아니다.

표 2

TP	FP	FN	TN
1	-1	-1	1
4	-2	-8	1
4	-4	-16	1
4	-16	-32	1
4	-16	-64	1
8	-4	-16	1
8	-4	-32	1
8	-8	-16	1
8	-8	-64	1
8	-16	-8	1
8	-16	-32	1
8	-16	-64	1
8	-32	-16	1
4	-64	-4	1
4	-64	-8	1
8	-96	-8	1

- [0075] [표 2]에서와 같이, 예시의 16개의 보상에서, TN의 보상은 1로 고정하고 기타 보상은 TN에 대한 비율을 나타내며, 기타 보상은 서로 동일하지 않도록 설정된다. 즉, 최적화된 보상 함수를 찾기 위해 TP, FP, FN 및 TN에 대해 서로 다른 보상 값이 할당된다.
- [0076] 한편, 본 발명의 실시예에서는, SOHO(the Solar and Heliospheric Observatory)/MDI(Michelson Doppler Imager)와 SDO(the Solar Dynamic Observatory)/HMI(Helioseismic and Magnetic Imager)의 00:00 UT(Universal Time, 세계시)의 태양 전면 자기장 지도(Full-disk Magnetogram)가 모델 훈련 및 테스트에서 사용된다. 여기서, SOHO/MDI와 SDO/HMI에 대해서는 잘 알려져 있으므로 구체적인 설명을 생략한다.
- [0077] SOHO의 과학 장비 중 하나인 MDI는 96분 주기로 픽셀당 1.98 arcsec로 1024×1024 전면 자기장 지도를 제공한다. SDO의 계측기 중 하나인 HMI는 MDI의 역할을 대신한다. HMI는 720초 케이던스에서 픽셀당 0.5 arcsec로 4096×4096 전면 자기장 지도를 생성한다.
- [0078] 전술한 MDI 데이터와 HMI 데이터를 병합하기 위해, MDI와 HMI의 공간 해상도가 블록 평균으로 512×512 크기로 조정된다. 또한, MDI와 HMI 사이에 잘 알려져 있는 관계식을 적용하여 MDI 데이터를 HMI 프로ksi 데이터로 변환한다. 데이터 값은 ±100 Gauss의 바이트 스케일로 표현되며, 태양 활성 영역의 자기장 특성을 잘 보여준다. 자기장 지도는 미 대기해양청의 GOES(Geostationary Operational Environmental Satellite) X선 플레어 데이터를 사용하여 중요 플레어 이벤트 일(≥M1.0 클래스) 또는 비중요 플레어 이벤트 일(<M1.0 클래스)로 별도로 레이블이 지정된다.
- [0079] 마지막으로, 훈련을 위해 훈련용 태양 주기 23 데이터(3914 데이터, MDI 1996년 5월~2008년 12월), 태양 주기 24 데이터(3492 데이터, 2009년 1월~2010년 12월 405 MDI, 2011년 1월~2019년 12월 3087 HMI)를 포함하는 데이터 세트를 구축한다. 태양 주기 23 및 태양 주기 24의 모든 기간을 사용하여, 훈련 및 테스트 기간 모두에서 태양 활동에 대한 태양 주기 효과를 완전히 고려한다. 데이터 세트의 예는 다음의 [표 3]에 설명되어 있는 바와 같으며, 이는 단지 예일 뿐이며, 이들로만 제한되는 것은 아니다.

표 3

	X	M	C	Non (<C)	Total
Training	93	647	1478	1696	3914
	740		3174		
Test	40	360	1240	1852	3492
	400		3092		

[0080]

- [0081] 한편, 모델의 성능은 다양한 관점에서 공정하게 평가되어야 하는데 이를 위해 4가지 기술 점수가 고려될 수 있다.

- [0082] 먼저, Heidke Skill Score(HSS) 기술 점수는 다음의 [수학식 6]과 같이 나타낼 수 있다.

- [0083] [수학식 6]

$$HSS = \frac{2[(TP \times TN) - (FP \times FN)]}{(TP + FN) \times (FN + TN) + (TP + FP) \times (FP + TN)}$$

[0084]

- [0085] 이러한 HSS는 랜덤 예측보다 개선 비율을 알려준다.

- [0086] 플레어 예측 평가에서 자주 사용되지만 그것은 이벤트 비율에 따라 범위가 달라질 수 있다.

- [0087] 불균형 데이터 세트에서 범위는 $-2(P \times N)/(P^2 + N^2) \sim 1$ 이며, 여기서 P는 양수 데이터의 개수이고, N은 음수 데이터의 개수이다. 음수 값은 랜덤 예측이 더 우수함을 의미하고, 0은 기술 예측이 없음을 의미하며, 1은 완벽한 예측을 나타낸다.

- [0088] 다음, 정밀도($\frac{TP}{TP+FP}$)와 재현율(recall)($\frac{TP}{TP+FN}$) 사이의 조화 평균인 F1 점수는 다음의 [수학식 7]과 같

이 주어질 수 있다.

[0089] [수학식 7]

$$F1 = \frac{TP}{TP + 0.5 \times (FP + FN)}$$

[0090]

[0091] 정밀도와 재현율이 둘 다 같을 때 평가 점수에 대해 동일한 기울기를 갖는다고 가정할 때 범위가 0에서 1인 모델 성능의 척도이다. F1 점수는 정밀도나 재현율에 편향되지 않으므로 예측 모델의 성능을 분석하는 데 좋은 평가 지표이다.

[0092] 다음, True Skill Statistic(TSS)은 -1에서 1 사이의 범위에서 모델을 평가한다. 이러한 TSS는 다음의 [수학식 8]과 같이 나타낼 수 있다.

[0093] [수학식 8]

$$TSS = \frac{TP}{TP + FN} - \frac{FP}{FP + TN}$$

[0094]

[0095] TSS는 플레어 모델을 비교하는 데 주로 사용된다. TSS는 테스트 세트에서 양수 값과 음수 값의 비율 변화에 영향을 받지 않기 때문에 플레어 모델을 비교하기 위한 표준 메트릭으로서 사용될 수 있다.

[0096] 다음, 숙련되지 않은 예측과 관련된 개선 사항을 알려주는 점을 제외하고, HSS와 유사한 Appleman's Skill Score(ApSS)는 다음의 [수학식 9]와 같이 산출될 수 있다.

[0097] [수학식 9]

$$ApSS = \begin{cases} \frac{TP-FP}{TP+FN}, & \text{if event rate} < 0.5 \\ \frac{TN-FN}{FP+TN}, & \text{if event rate} \geq 0.5 \end{cases}$$

[0098]

[0099] 여기서, 이벤트 비율(event rate)은 전체 데이터 개수에 대한 플레어 이벤트 개수의 비율이다. ApSS는 FN과 FP를 동등하게 고려하여 모델의 성능에 대한 좋은 개요를 제공할 수 있다. 그 규모는 $\frac{N}{P}$ 에서 1까지의 데이터 분포에 따라 다르다. ApSS는 불균형 데이터 분류의 일반적인 결과인 FP가 TP보다 클 때 음수가 되기 때문에 까다로운 평가 메트릭이다. 음의 ApSS를 가지면서 높은 TSS를 달성하는 것은 주요 태양 플레어 모델에서 일반적이다.

[0100] 본 발명의 실시예에서는 전술한 데이터 세트(표 2 참조)를 사용하여 학습되는 전술한 본 발명의 실시예에 따른 DQN(또는 더블 DQN)(100) 모델에 대해 전술한 HSS, F1, TSS, Apss의 네 가지 기술 점수를 사용하여 모델의 성능을 분석한다.

[0101] 전술한 바와 같이, 다른 기술 점수는 모델 성능의 다른 측면을 나타낸다. 모델 성능에 대한 강화 학습의 효과를 측정하기 위해 각 모델에 대한 "최적 기술 점수"를 제시한다. "최적 기술 점수"는 각 기술 점수에 최적화된 모델에서 측정된 스킬 점수이다.

[0102] 모델을 최적화하기 위해 모델에 두 가지 방법이 적용될 수 있다. 먼저, 과적합 문제를 피하고 최상의 성능을 찾기 위해 심층 강화학습 모델을 충분한 에포크(epoch) 동안 훈련하고 에포크에서 가장 높은 점수를 받은 모델이 선택된다. 둘째, 최적의 기술 점수를 측정하기 위한 예측 임계값을 결정한다. 이러한 전략은 특정 작업에 대한 적절한 모델 개발 옵션을 제안할 수 있다.

[0103] 본 발명의 실시예에 따른 태양 중요 플레어 예측 모델의 성능 비교를 위해 도 2에 도시된 바와 같은, CNN 모델이 250 에포크 내에서 최적화된다. 그러나, 강화학습 모델은 훈련 시간 때문에 100 에포크 동안 훈련된다. Nvidia RTX 2080 GPU에서 250 에포크 CNN 모델에 대한 훈련 시간은 4시간인 반면 100 에포크 강화학습 모델에 대한 훈련 시간은 그보다 긴 28시간이다. 250 에포크와 100 에포크 내에서 최적화된 CNN 모델의 성능을 비교한 결과 전자가 후자를 능가하는 것으로 나타났다. 또한, CNN 모델이 250 에포크 이후에 과대적합된 것을 발견할 수 있었다. 결과를 더 엄격하게 비교하기 위해 CNN 모델은 250 에포크 동안 훈련되고 강화학습 모델은 100 에포크 동안 훈련된다.

[0104] 본 발명의 실시예에 따른 태양 중요 플레어 예측 모델을 결정하기 위해 [표 1]에 나타난 바와 같은 16개의 보상

함수, 2개의 강화 학습 방법(DQN 및 더블 DQN) 및 2개의 목표 에이전트 업데이트 시간(1개의 에포크 및 2개의 에포크)을 기반으로 64개의 심층 강화학습 모델을 만든다.

[0105] 전술한 16개의 보상함수에서, TN의 보상은 1로 고정하고 기타 보상은 TN에 대한 비율을 나타낸다. 최적화된 보상 함수를 찾기 위해 TP, FP, FN 및 TN에 대해 서로 다른 보상 값을 할당한다. 희귀 이벤트의 비용을 늘리기 위해 많은 보상 함수가 FP보다 FN에 더 높은 페널티를 가진다.

[0106] 다음의 [표 4]는 각각 HSS, F1, TSS 및 ApSS에 대해 최적화된 태양 중요 플레이어 예측 모델의 결과를 나타낸다. 64개 모델 중 가장 높은 성능과 두 번째로 높은 성능은 HSS가 0.44와 0.43, F1이 0.52와 0.51, TSS가 0.59와 0.58, ApSS가 0.12와 0.11이다.

표 4

모델 번호	보상 함수				방법	업데이트 시간	최적 점수			
	TP	FP	FN	TN			HSS	F1	TSS	ApSS
0					CNN		0.38	0.47	0.53	0.09
HSS 최적 모델										
1	8	-64	-8	1	DQN	2 epoch	0.44	0.51	0.59	0.10
2	8	-16	-64	1	Double DQN	1 epoch	0.44	0.52	0.55	0.08
3	4	-16	-64	1	Double DQN	1 epoch	0.43	0.50	0.58	0.11
4	8	-8	-64	1	DQN	1 epoch	0.43	0.51	0.57	0.09
5	8	-96	-8	1	Double DQN	2 epoch	0.43	0.50	0.51	0.10
F1 최적 모델										
2	8	-16	-64	1	Double DQN	1 epoch	0.44	0.52	0.55	0.08
1	8	-64	-8	1	DQN	2 epoch	0.44	0.51	0.59	0.10
4	8	-8	-64	1	DQN	1 epoch	0.43	0.51	0.57	0.09
TSS 최적 모델										
1	8	-64	-8	1	DQN	2 epoch	0.44	0.51	0.59	0.10
3	4	-16	-64	1	Double DQN	1 epoch	0.43	0.50	0.58	0.11
6	8	-64	-8	1	Double DQN	2 epoch	0.43	0.49	0.58	0.10
7	4	-16	-64	1	DQN	1 epoch	0.39	0.47	0.58	0.09
ApSS 최적 모델										
8	4	-4	-16	1	Double DQN	1 epoch	0.41	0.49	0.56	0.12
9	8	-96	-8	1	Double DQN	1 epoch	0.39	0.47	0.48	0.12
10	8	-4	-16	1	DQN	2 epoch	0.40	0.47	0.54	0.12
11	4	-2	-8	1	Double DQN	2 epoch	0.39	0.47	0.49	0.12
12	4	-4	-16	1	Double DQN	2 epoch	0.40	0.46	0.47	0.12
13	8	-4	-16	1	Double DQN	2 epoch	0.38	0.44	0.45	0.12
14	8	-4	-32	1	Double DQN	2 epoch	0.41	0.49	0.53	0.12
15	8	-8	-64	1	Double DQN	2 epoch	0.40	0.47	0.50	0.12

[0107]

[0108] HSS, F1 또는 TSS에서 가장 높은 점수 또는 두 번째로 높은 점수를 얻은 모델의 보상 함수가 [표 2]에 나타난 보상 함수이다. 그러나, ApSS에서는 많은 모델이 가장 높은 ApSS에 대해 동점을 기록했기 때문에 가장 높은 점수를 얻은 모델을 제시한다.

[0109] 각각의 평가 메트릭에 대해 나열된 모델 개수는 모델 분석에 영향을 미치지 않음을 알 수 있다. 보상 함수, 방법, 업데이트 시간 등을 고려하여 어떤 모델이 더 낫다고 평가하기는 어려우므로, HSS, F1, TSS에서 최고 점수와 두 번째로 높은 점수를 달성한 모든 모델과 ApSS에서 최고 점수를 달성한 모델을 모두 제시할 수 있다.

[0110] [표 4]에서 모델 번호 0은 다른 심층 강화학습 모델과 동일한 구조를 가진 CNN 플레이어 모델이다. 심층 강화학습 모델과의 비교를 위해 CNN 모델 결과가 제시된 것이다.

[0111] 태양 중요 플레이어 예측 모델은 HSS에서 0.38에서 0.44로, F1에서 0.47에서 0.52로, TSS에서 0.50에서 0.58로, ApSS에서 0.09에서 0.12로 증가하여 CNN 모델보다 더 나은 점수를 보여주는 것을 알 수 있다.

- [0112] 모델 번호 1, 2, 3 및 4는 HSS, F1 및/또는 TSS에서 좋은 성능을 보여주는 것을 알 수 있다.
- [0113] HSS 및 TSS에서 모델은 네 가지 모델 설정 모두에서 최적화된 성능을 달성하고 있다. HSS 최적화 모델은 서로 다른 5개의 보상 함수로 좋은 성능을 보이는 반면, TSS 최적화 모델은 두 가지 보상 함수만으로 좋은 성능을 보이고 있다. HSS, F1, TSS에서 좋은 성능을 보인 모든 모델은 보상 함수에서 FP 또는 FN에서 -64 이하의 보상을 포함하고 있다. 이는 높은 HSS, F1 또는 TSS를 달성하기 위해 모델이 잘못된 예측에 대해 강력한 패널티를 요구함을 나타낸다.
- [0114] ApSS에서는 더블 DQN을 사용하고 업데이트 시간이 2 에포크인 많은 모델이 최상의 결과를 얻고 있음을 알 수 있다. ApSS 최적화 모델은 HSS에서 CNN 모델보다 비슷하거나 약간 더 나은 최적 점수를 보이지만 대부분은 CNN 모델보다 TSS가 낮다.
- [0115] 태양 중요 플레이어 예측 모델의 불확실성을 추정하기 위해, 모델 번호 3을 서로 다른 랜덤 시드로 10회 훈련시키고 4가지 기술 점수의 평균과 표준 편차를 계산하면, 그 결과는 HSS의 경우 0.42 ± 0.02 , F1의 경우 0.49 ± 0.01 , TSS의 경우 0.58 ± 0.02 , ApSS의 경우 0.10 ± 0.01 이다.
- [0116] 전술한 강화학습에서, 훈련 진행 상황을 공정하게 평가하는 것은 어려운 일이다. 평균 총 보상의 메트릭은 에이전트의 가중치에 대한 작은 조정으로 인해 에이전트의 출력에 상당한 변화가 발생하기 때문에 고도로 분포될 수 있다. 예를 들어, 도 4는 테스트 데이터 세트를 사용하여 모델 번호 3을 기반으로 10개 모델의 평균 보상을 추적하는 훈련 곡선을 나타낸 것으로, 불확실성을 추정하는 데 사용된 10개 모델의 평균 훈련 성능을 제공한다. 곡선에 노이즈가 있지만 훈련이 진행됨에 따라 곡선의 분포가 감소하고 높은 값으로 수렴됨을 알 수 있다.
- [0117] 전술한 바와 같이, 본 발명의 실시예에 따른 딥러닝 강화학습을 이용한 태양 중요 플레이어 예측 방법에 따르면, HSS, F1, TSS 및 ApSS와 같은 우수한 기술 점수로 주요 태양 플레이어를 성공적으로 예측할 수 있고, 또한 중요(강한) 플레이어($\geq M$ -class)가 중요하지 않은(약한) 플레이어($\leq C$ -class)보다 훨씬 더 가치가 있고 약한 플레이어보다 강한 플레이어에 대해 더 높은 보상과 더 강한 패널티를 할당한다면 심층 강화학습 플레이어 모델은 약한 것보다 강한 것을 더 중요하게 고려함으로써 최적화될 수 있으며, 또한 모델의 성능은 보상 함수, 학습 방법 및 타깃 에이전트 업데이트 시간, 예를 들어, 1 에포크 및 2 에포크에 따라 달라질 수 있고, 또한, HSS는 $0.38(\text{CNN}) \sim 0.44(\text{본 발명})$, F1은 $0.47 \sim 0.52$, TSS는 $0.53 \sim 0.59$, TSS는 $0.09 \sim 0.12$ 로 동일한 구조의 CNN 모델보다 성능이 훨씬 뛰어남을 알 수 있다.
- [0118] 이하, 전술한 본 발명의 실시예에 따른 딥러닝 강화학습을 이용한 태양 중요 플레이어 예측 방법을 구현하는 장치에 대해 설명한다.
- [0119] 도 5는 본 발명의 실시예에 따른 딥러닝 강화학습을 이용한 태양 중요 플레이어 예측 장치의 구성 블록도이다.
- [0120] 도 5에 도시된 바와 같이, 본 발명의 실시예에 따른 태양 중요 플레이어 예측 장치(300)는 데이터 수신부(310), 학습부(320) 및 예측부(330)를 포함한다. 도 5에는 본 발명의 실시예와 관련이 있는 구성요소들만이 도시되어 있다. 따라서, 본 발명이 속하는 기술 분야에서 통상의 기술자라면 도 5에 도시된 구성요소들 외에 다른 범용적인 구성요소들이 더 포함될 수 있음을 알 수 있다.
- [0121] 데이터 수신부(310)는 태양 중요 플레이어 예측 모델(340)의 학습 및 태양 중요 플레이어 예측에 사용될 데이터를 소정의 인터페이스를 통하여 수신할 수 있다. 실시예에 따라, 데이터 수신부(310)는 태양 중요 플레이어 예측 모델(340)의 학습에 사용될 학습 데이터와 태양 중요 플레이어 예측에 사용될 입력 데이터를 구분하여 수신할 수 있다. 여기서, 학습 데이터에는 태양 중요 플레이어 예측 모델(340)의 검증에 사용되는 테스트 데이터가 포함되어 있다. 데이터 수신부(310)는 학습 데이터에서 소정의 기준에 따라 테스트 데이터를 분리할 수 있다. 예컨대, 수신된 복수의 태양 자기장 지도들을 시간 순서에 따라 일정 비율로 분할함으로써 학습 데이터에서 테스트 데이터를 분리할 수 있다.
- [0122] 또한, 데이터 수신부(310)는 각각이 수신된 데이터에 포함된 태양 전면 자기장 지도들 각각에 대응하는 태양 플레이어 발생 정보와 태양 플레이어의 발생 여부를 예측하기 위한 예측 대상 데이터(즉, 예측 대상 태양 전면 자기장 지도)를 수신할 수도 있다.
- [0123] 학습부(320)는 데이터 수신부(310)에 의해 수신되거나 데이터 수신부(310)에 의해 생성된 학습 데이터 및/또는 태양 플레이어 발생 정보를 이용하여 본 발명의 실시예에 따른 DQN(CNN 포함)(도 2, 도 3 참조)을 학습함으로써 태양 중요 플레이어 예측 모델(340)을 생성할 수 있다. 예컨대, 학습부(320)는 태양 전면 자기장 지도와 이에 대응하는 태양 플레이어 발생 정보를 사용하여 DQN(CNN 포함)의 입력과 출력으로 사용하여 학습함으로써 태양 중요

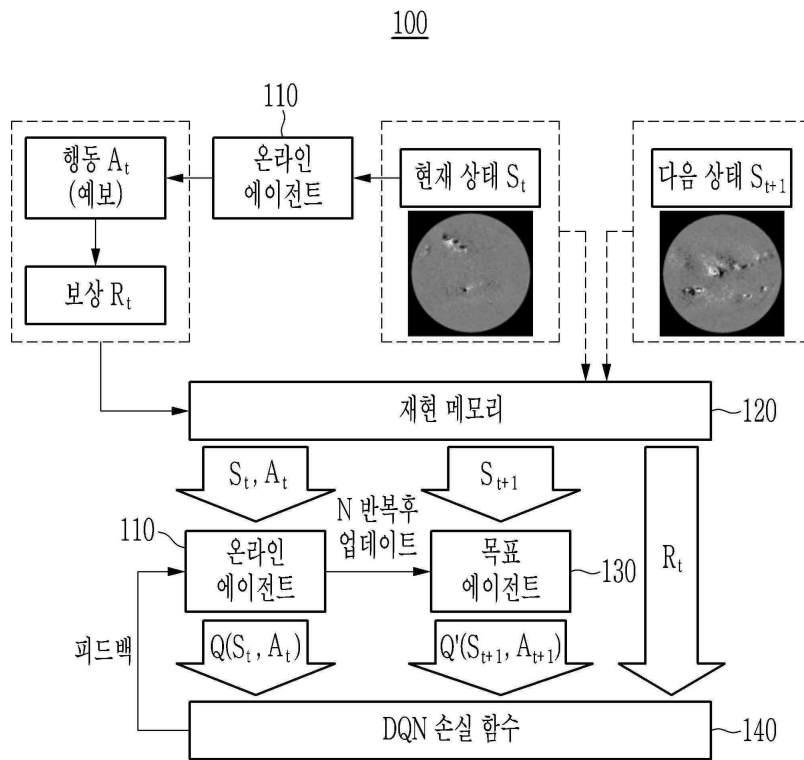
플레이어 예측 모델(340)을 생성할 수 있다.

- [0124] 또한, 학습부(320)는 생성된 태양 중요 플레이어 예측 모델(340)을 검증(테스트)할 수 있다.
- [0125] 구체적으로, 학습부(320)는 데이터 수신부(310)에 의해 수신되거나 데이터 수신부(310)에 의해 생성된 테스트 데이터를 이용하여 태양 중요 플레이어 예측 모델(340)을 검증할 수 있다. 구체적으로, 학습부(320)는 테스트 데이터에 포함된 각 태양 전면 자기장 지도를 학습된 태양 중요 플레이어 예측 모델에 입력하였을 때의 출력과 입력된 태양 전면 자기장 지도에 대응하는 태양 플레이어 발생 정보를 사용함으로써 학습된 태양 중요 플레이어 예측 모델(340)을 검증할 수 있다.
- [0126] 예측부(330)는 태양 중요 플레이어 예측 모델(340)을 이용하여 데이터 수신부(310)에 의해 수신된 입력 데이터(예측 대상 태양 전면 자기장 지도)에 대한 태양 중요 플레이어 예측 결과를 출력함으로써 태양 중요 플레이어의 발생 여부를 예측할 수 있다.
- [0127] 예컨대, 예측부(330)는 예측 대상 데이터가 데이터 수신부(310)로부터 수신된 경우, 태양 중요 플레이어 예측 모델(340)에 예측 대상 데이터를 입력하고 그에 따른 결과를 출력함으로써 태양 중요 플레이어의 발생 여부를 예측할 수 있다.
- [0128] 이하, 본 발명의 실시예에 따른 태양 중요 플레이어 예측 방법에 대해 설명한다.
- [0129] 도 6은 본 발명의 실시예에 따른 태양 중요 플레이어 예측 방법의 개략적인 흐름도이다. 도 6에 도시된 본 발명의 실시예에 따른 태양 중요 플레이어 예측 방법은 전술한 도 5를 참조하여 설명한 태양 중요 플레이어 예측 장치(300)에 의해 수행될 수 있다.
- [0130] 도 5 및 도 6을 참조하면, 태양 중요 플레이어 예측 장치(300)의 데이터 수신부(310)가 태양 중요 플레이어 예측 모델(340)의 학습 및 검증과 태양 중요 플레이어 예측에 사용되는 데이터를 수신할 수 있다(S110).
- [0131] 다음, 태양 중요 플레이어 예측 장치(300)의 학습부(320)는 수신된 데이터에 포함된 학습 데이터를 이용하여 본 발명의 실시예에 따른 DQN(CNN 포함)(도 2, 도 3 참조)을 학습함으로써 태양 중요 플레이어 예측 모델(340)을 생성할 수 있다(S120).
- [0132] 또한, 태양 중요 플레이어 예측 장치(300)의 학습부(320)는 수신된 데이터에 포함된 테스트 데이터를 이용하여 태양 중요 플레이어 예측 모델(340)을 검증할 수 있다(S130).
- [0133] 마지막으로, 태양 중요 플레이어 예측 장치(300)의 예측부(330)는 수신된 예측 대상 태양 전면 자기장 지도를 태양 중요 플레이어 예측 모델(340)에 입력함으로써 태양 중요 플레이어 발생 여부를 예측할 수 있다(S140).
- [0134] 다음, 본 발명의 다른 실시예에 따른 태양 중요 플레이어 예측 장치에 대해 설명한다.
- [0135] 도 7은 본 발명의 다른 실시예에 따른 태양 중요 플레이어 예측 장치(400)의 구성 블록도이다.
- [0136] 도 7에 도시된 바와 같이, 태양 중요 플레이어 예측 장치(400)는 중앙 처리 장치(Central Processing Unit, CPU)(410), 작업 메모리(420), 입출력 장치(430), 저장 장치(440) 및 시스템 버스(450)를 포함한다. 여기서, 태양 중요 플레이어 예측 장치(400)는 태양 중요 플레이어 예측을 수행하기 위한 전용 장치로서 제공될 수도 있지만, 다양한 프로그램을 구동하기 위한 범용 컴퓨터로 구현될 수도 있다.
- [0137] CPU(410)는 태양 중요 플레이어 예측 장치(400)에서 수행될 소프트웨어(응용 프로그램, 운영 체제, 장치 드라이버 등)를 실행한다. CPU(410)는 작업 메모리(420)에 로드되는 운영 체제(OS, 미도시)를 실행할 수 있다. CPU(410)는 운영 체제(OS) 기반에서 구동될 다양한 응용 프로그램들(Application Program)이나 예측 모델 등을 실행할 것이다.
- [0138] 작업 메모리(420)에는 운영 체제(OS)나 응용 프로그램들이 로드될 것이다. 태양 중요 플레이어 예측 장치(400)의 부팅시에 저장 장치(440)에 저장된 OS 이미지(미도시)가 부팅 시퀀스에 의거하여 작업 메모리(420)로 로드될 것이다.
- [0139] 운영 체제(OS)에 의해서 태양 중요 플레이어 예측 장치(400)의 제반 입출력 동작들이 지원될 수 있다. 마찬가지로, 사용자에 의하여 선택되거나 기본적인 서비스 제공을 위해서 응용 프로그램들이 작업 메모리(420)에 로드될 수 있다. 작업 메모리(420)는 SRAM(Static Random Access Memory)이나 DRAM(Dynamic Random Access Memory)과 같은 휘발성 메모리이거나, PRAM, MRAM, ReRAM, FRAM, NOR 플래시 메모리 등과 같은 비휘발성 메모리일 수 있다.

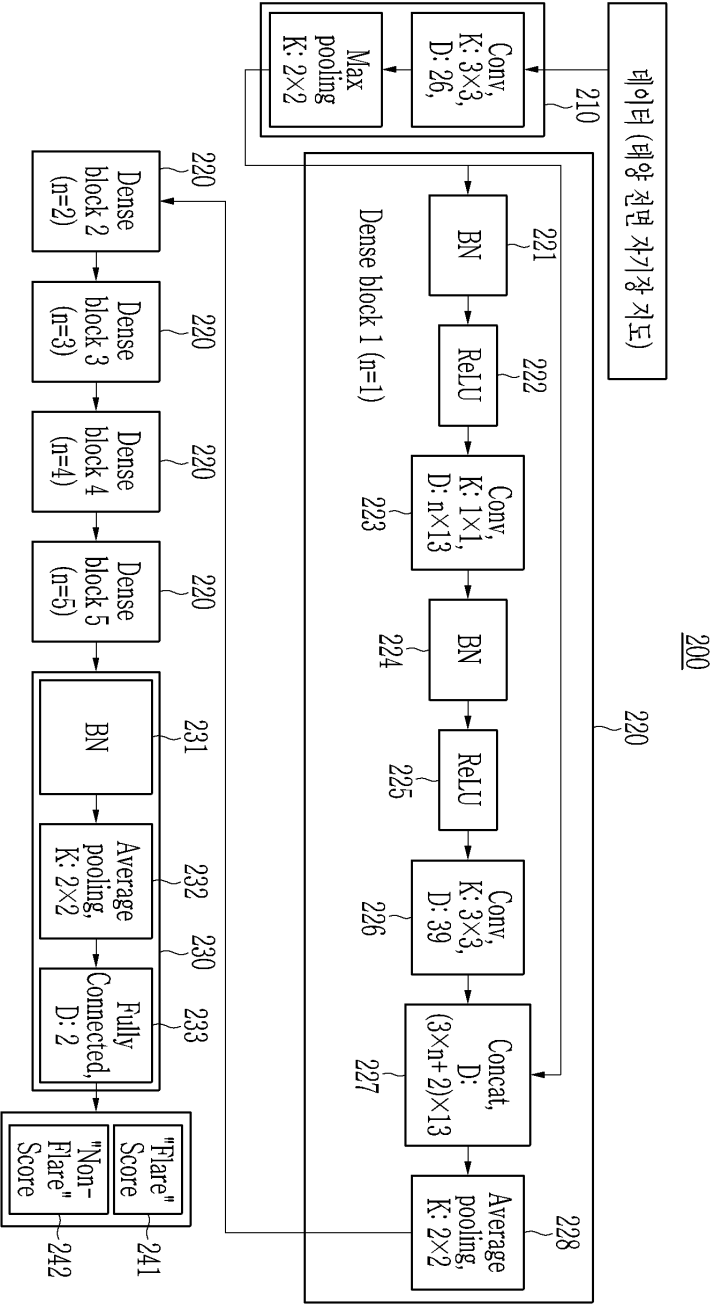
- [0140] 입출력 장치(430)는 사용자 인터페이스 장치들로부터의 사용자 입력 및 출력을 제어한다. 예컨대, 입출력 장치(430)는 키보드, 마우스, 터치 패드와 같은 입력 장치를 구비하여 태양 중요 플레어 발생 예측에 필요한 파일이나 태양 중요 플레어 예측 모델(340) 생성에 필요한 구성 정보를 입력받을 수 있다. 그리고, 입출력 장치(430)는 모니터 등의 출력 장치를 구비하여 태양 중요 플레어 예측 장치(400)의 태양 중요 플레어 예측 모델(340) 생성, 태양 중요 플레어 예측 모델 검증, 및 태양 중요 플레어 예측 동작에서의 경과 및 처리 결과 등을 표시할 수 있다.
- [0141] 저장 장치(440)는 태양 중요 플레어 예측 장치(400)의 저장 매체(Storage Medium)로서 제공된다. 저장 장치(440)는 응용 프로그램들, 운영 체제 이미지(OS Image) 및 각종 데이터를 저장할 수 있다. 또한, 저장 장치(440)는 태양 중요 플레어 예측 모델(340)의 학습 및 검증, 그리고 태양 중요 플레어 예측 모듈(340)의 실행에 필요한 데이터 및 이들의 실행에 의해 생성된 태양 중요 플레어 예측 모델(340)이나 생성된 데이터 등을 저장할 수 있다. 저장 장치(440)는 메모리 카드(MMC, eMMC, SD, MicroSD 등)나 하드 디스크 드라이브(HDD)로 제공될 수도 있다. 저장 장치(440)는 대용량의 저장 능력을 갖는 낸드 플래시 메모리(NAND-type Flash memory)를 포함할 수 있다. 또한, 저장 장치(440)는 FRAM, MRAM, ReRAM, FRAM 등의 차세대 불휘발성 메모리나 NOR 플래시 메모리를 포함할 수도 있다.
- [0142] 시스템 버스(450)는 태양 중요 플레어 예측 장치(400)의 내부에서 네트워크를 제공하기 위한 인터커넥터로 제공될 수 있다. 시스템 버스(450)를 통해서 CPU(410), 작업 메모리(420), 입출력 장치(430) 및 저장 장치(440)가 전기적으로 연결되고 상호 데이터를 교환할 수 있다. 하지만, 시스템 버스(450)의 구성은 전술한 설명에만 국한되지 않으며, 효율적인 관리를 위한 중재 수단들을 더 포함할 수도 있다.
- [0143] 본 발명의 실시예에서 사용되는 구성요소 또는 “~부”는 메모리 상의 소정 영역에서 수행되는 태스크, 클래스, 서브 루틴, 프로세스, 오브젝트, 실행 쓰레드, 프로그램과 같은 소프트웨어(software)나, FPGA(field-programmable gate array)나 ASIC(application-specific integrated circuit)과 같은 하드웨어(hardware)로 구현될 수 있으며, 또한 상기 소프트웨어 및 하드웨어의 조합으로 이루어질 수도 있다. 상기 구성요소 또는 ‘~부’ 등은 컴퓨터로 판독 가능한 저장 매체에 포함되어 있을 수도 있고, 복수의 컴퓨터에 그 일부가 분산되어 분포될 수도 있다.
- [0144] 이상에서 설명한 본 발명의 실시예는 장치 및 방법을 통해서만 구현이 되는 것은 아니며, 본 발명의 실시예의 구성에 대응하는 기능을 실현하는 프로그램 또는 그 프로그램이 기록된 기록 매체를 통해 구현될 수도 있다.
- [0145] 이상에서 본 발명의 실시예에 대하여 상세하게 설명하였지만 본 발명의 권리범위는 이에 한정되는 것은 아니고 다음의 청구범위에서 정의하고 있는 본 발명의 기본 개념을 이용한 당업자의 여러 변형 및 개량 형태 또한 본 발명의 권리범위에 속하는 것이다.

도면

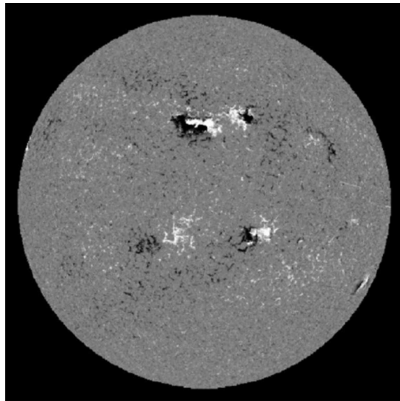
도면1



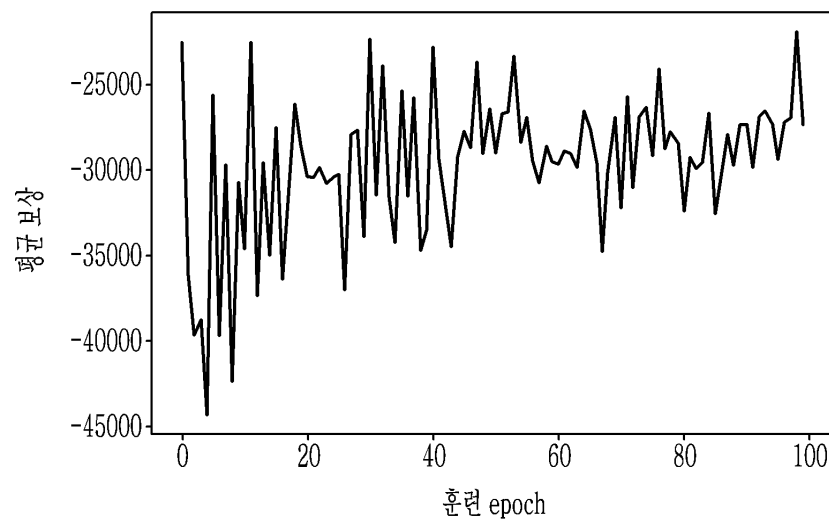
도면2



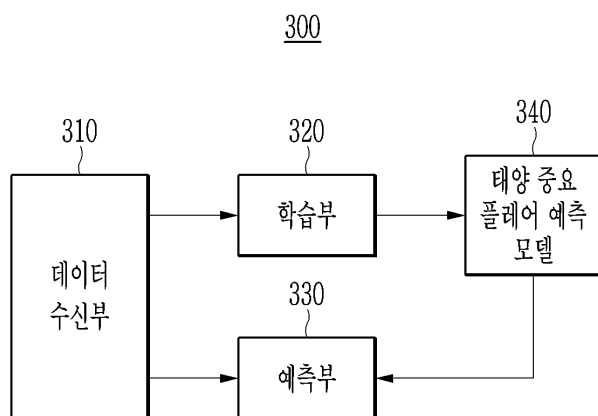
도면3



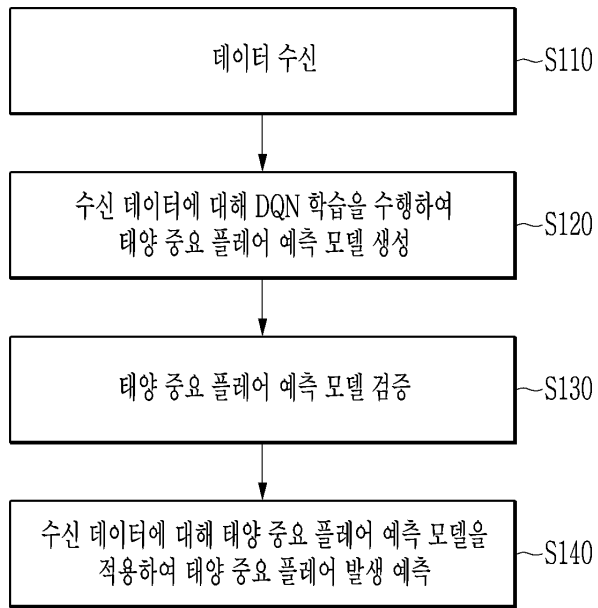
도면4



도면5



도면6



도면7

