(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2018-0091499
(43) 공개일자 2018년08월16일

(51) 국제특허분류(Int. Cl.)
G01W 1/10 (2006.01) *G06N 5/00* (2006.01)
(52) CPC특허분류
G01W 1/10 (2013.01)
G06N 5/00 (2013.01)
(21) 출원번호 10-2017-0016763
(22) 출원일자 2017년02월07일
심사청구일자 2017년02월07일

(71) 출원인
울산과학기술원
울산광역시 울주군 언양읍 유니스트길 50
한국전자통신연구원
대전광역시 유성구 가정로 218 (가정동)
(72) 발명자
임정호
울산광역시 울주군 언양읍 유니스트길 50
김미애
울산광역시 울주군 언양읍 유니스트길 50
(뒷면에 계속)
(74) 대리인
특허법인 무한

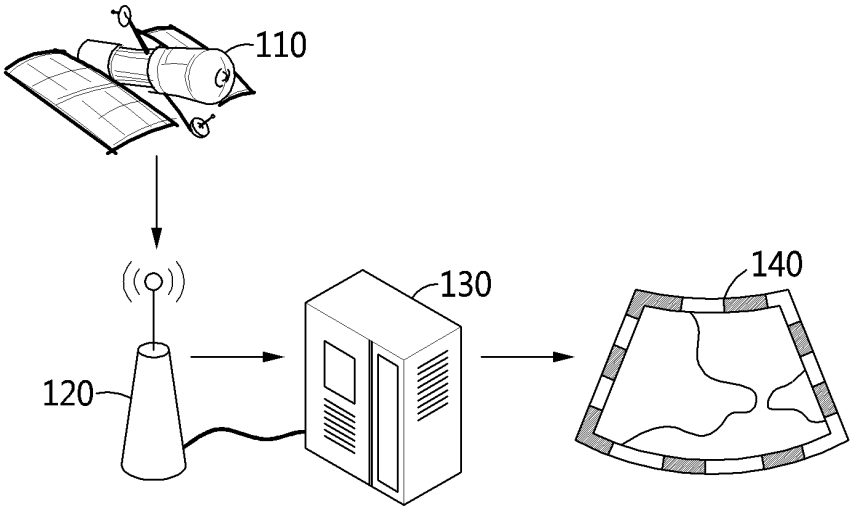
전체 청구항 수 : 총 21 항

(54) 발명의 명칭 **성층권 침투 대류운 출현 지역의 분류 장치 및 방법**

(57) 요 약

위성자료에서 성층권 침투 대류운의 출현 및 비출현의 기준이 되는 지역을 탐색하는 단계, 기준이 되는 지역에 대한 정보로부터 성층권 침투 대류운의 출현 및 비출현을 판단하기 위한 입력 변수를 추출하는 단계 및 입력 변수를 적용하여 학습기법을 이용한 트레이닝을 통해 성층권 침투 대류운의 출현 및 비출현을 분류하는 분류 모델을 구축하는 단계를 포함하는 분류 모델 구축 방법이 제공될 수 있다.

대 표 도 - 도1



(52) CPC특허분류
G01W 2201/00 (2013.01)
(72) 발명자
정원찬
대전광역시 유성구 가정로 218
김지인
대전광역시 유성구 가정로 218

명 세 서

청구범위

청구항 1

위성자료에서 성층권 침투 대류운의 출현 및 비출현의 기준이 되는 지역을 탐색하는 단계;

상기 기준이 되는 지역에 대한 정보로부터 상기 성층권 침투 대류운의 출현 및 비출현을 판단하기 위한 입력 변수를 추출하는 단계; 및

상기 입력 변수를 적용하여 학습기법을 이용한 트레이닝을 통해 상기 성층권 침투 대류운의 출현 및 비출현의 지역을 분류하는 분류 모델을 구축하는 단계

를 포함하는 분류 모델 구축 방법.

청구항 2

제1항에 있어서,

상기 입력 변수를 제1 입력 변수 세트 및 제2 입력 변수 세트로 임의로 분류하는 단계; 및

상기 제2 입력 변수 세트를 적용하여 상기 트레이닝을 통해 구축되는 상기 분류 모델을 테스트하는 단계

를 더 포함하고,

상기 분류 모델을 구축하는 단계에서 상기 트레이닝에 적용되는 상기 입력 변수는 상기 제1 입력 변수 세트에 대응하는 분류 모델 구축 방법.

청구항 3

제1항에 있어서,

상기 위성자료를 관측 시점이 서로 다른 제1 위성자료 및 제2 위성자료로 구분하는 단계;

상기 제2 위성자료에서 실제 성층권 침투 대류운의 출현 및 비출현 지역을 지정하는 단계;

상기 트레이닝을 통해 구축되는 상기 분류 모델을 상기 제2 위성자료에 적용하여 상기 성층권 침투 대류운의 출현 및 비출현 지역을 분류하는 단계; 및

상기 지정하는 단계의 상기 실제 성층권 침투 대류운의 출현 및 비출현 지역과 상기 분류하는 단계의 상기 성층권 침투 대류운의 출현 및 비출현 지역을 비교하여 상기 분류 모델을 검증하는 단계

를 더 포함하고,

상기 탐색하는 단계에서 탐색하는 상기 위성자료는 상기 제1 위성자료에 대응하는 분류 모델 구축 방법.

청구항 4

제1항에 있어서,

상기 탐색하는 단계는 상기 위성자료가 위성에서 수신되는 가시영상 및 수치일기예보 대류권계면 온도자료에 해당하는 분류 모델 구축 방법.

청구항 5

제4항에 있어서,

상기 가시영상의 광휘값을 반사도로 변환하는 단계

를 더 포함하는 분류 모델 구축 방법.

청구항 6

제1항에 있어서,

상기 추출하는 단계는 위성에서 수신되는 적외영상의 데이터로부터 입력 변수를 추출하는 분류 모델 구축 방법.

청구항 7

제6항에 있어서,

상기 적외영상의 광휘값을 밝기온도로 변환하는 단계

를 더 포함하는 분류 모델 구축 방법.

청구항 8

제1항에 있어서,

상기 학습기법은 DT(Decision Trees), RF(Random Forest), ERT(Extremely Randomized Trees), 및 SVM(Support Vector Machines)을 포함하는 분류 모델 구축 방법.

청구항 9

제2항에 있어서,

상기 테스트는 통계치를 이용한 성능 평가를 통해 수행되는 분류 모델 구축 방법.

청구항 10

제1항에 있어서,

상기 추출하는 단계는 상기 입력 변수를 윈도 사이즈 별로 추출하는 분류 모델 구축 방법.

청구항 11

위성자료에서 성층권 침투 대류운의 출현 및 비출현의 기준이 되는 지역을 탐색하는 탐색부;

상기 기준이 되는 지역에 대한 정보로부터 상기 성층권 침투 대류운의 출현 및 비출현을 판단하기 위한 입력 변수를 추출하는 변수 추출부; 및

상기 입력 변수를 적용하여 학습기법을 이용한 트레이닝을 통해 상기 성층권 침투 대류운의 출현 및 비출현의 지역을 분류하는 분류 모델을 구축하는 모델 생성부

를 포함하는 분류 모델 구축 장치.

청구항 12

제11항에 있어서,

상기 입력 변수를 제1 입력 변수 세트 및 제2 입력 변수 세트로 임의로 분류하는 변수 분류부; 및

상기 제2 입력 변수 세트를 적용하여 상기 트레이닝을 통해 구축되는 상기 분류 모델을 테스트하는 테스트부

를 더 포함하고,

상기 모델 생성부가 상기 트레이닝에 적용하는 상기 입력 변수는 상기 제1 입력 변수 세트에 대응하는 분류 모델 구축 장치.

청구항 13

제11항에 있어서,

상기 위성자료를 관측 시점이 서로 다른 제1 위성자료 및 제2 위성자료로 구분하여 분류하는 위성자료 분류부;

상기 제2 위성자료에서 실제 성층권 침투 대류운의 출현 및 비출현 지역을 지정하는 지역 지정부;

상기 트레이닝을 통해 구축되는 상기 분류 모델을 상기 제2 위성자료에 적용하여 상기 성층권 침투 대류운의 출현 및 비출현 지역을 분류하는 지역 분류부; 및

상기 지정부의 상기 실제 성층권 침투 대류운의 출현 및 비출현 지역과 상기 지역 분류부의 상기 성층권 침투 대류운의 출현 및 비출현 지역을 비교하여 상기 분류 모델을 검증하는 검증부

를 더 포함하고,

상기 탐색부가 탐색하는 상기 위성자료는 상기 제1 위성자료에 대응하는 분류 모델 구축 장치.

청구항 14

제11항에 있어서,

상기 탐색부는 상기 위성자료가 위성에서 수신되는 가시영상 및 수치일기예보 대류권계면 온도자료에 해당하는 분류 모델 구축 장치.

청구항 15

제14항에 있어서,

상기 가시영상의 광휘값을 반사도로 변환하는 제1 변환부를

를 더 포함하는 분류 모델 구축 장치.

청구항 16

제11항에 있어서,

상기 변수 추출부는 위성에서 수신되는 적외영상의 데이터로부터 입력 변수를 추출하는 분류 모델 구축 장치.

청구항 17

제16항에 있어서,

상기 적외영상의 광휘값을 밝기온도로 변환하는 제2 변환부

를 더 포함하는 분류 모델 구축 장치.

청구항 18

제11항에 있어서,

상기 학습기법은 DT(Decision Trees), RF(Random Forest), ERT(Extremely Randomized Trees) 및 SVM(Support Vector Machines)을 포함하는 분류 모델 구축 장치.

청구항 19

제12항에 있어서,

상기 테스트는 통계치를 이용한 성능 평가를 통해 수행되는 분류 모델 구축 장치.

청구항 20

제11항에 있어서,

상기 변수 추출부는 상기 입력 변수를 윈도 사이즈 별로 추출하는 분류 모델 구축 장치.

청구항 21

제1항 내지 제10항 중 어느 한 항의 분류 모델 구축 방법을 수행하는 프로그램을 수록한 컴퓨터 판독 가능 기록 매체.

발명의 설명

기 술 분 야

[0001] 성층권 침투 대류운 출현 지역을 분류하는 장치 및 방법에 연관되며, 보다 특정하게는 기계학습기법을 이용하여

분류하는 장치 및 방법에 연관된다.

배 경 기 술

[0002] 정지 궤도 기상 위성의 관측 자료를 이용하여 데이터베이스 처리를 통해 기상 현상을 분석하거나, 위성의 영상 자료를 분석 및 처리하여 일기 예보 또는 날씨 데이터에 반영하는 것은 일반적이다.

[0003] 기계학습(Machine Learning)이란 인공지능의 한 분야로, 정보의 학습 작업을 반복적으로 수행함으로써, 빅데이터를 분석해 이후를 예측하는 응용이 가능하다.

발명의 내용

해결하려는 과제

과제의 해결 수단

[0004] 일측에 따르면, 위성자료에서 성층권 침투 대류운의 출현 및 비출현의 기준이 되는 지역을 탐색하는 단계, 기준이 되는 지역에 대한 정보로부터 성층권 침투 대류운의 출현 및 비출현을 판단하기 위한 입력 변수를 추출하는 단계 및 입력 변수를 적용하여 학습기법을 이용한 트레이닝을 통해 성층권 침투 대류운의 출현 및 비출현의 지역을 분류하는 분류 모델을 구축하는 단계를 포함하는 분류 모델 구축 방법이 제공될 수 있다.

[0005] 일실시예에 따르면, 분류 모델 구축 방법은 입력 변수를 제1 입력 변수 세트 및 제2 입력 변수 세트로 임의로 분류하는 단계 및 제2 입력 변수 세트를 적용하여 트레이닝을 통해 구축되는 분류 모델을 테스트하는 단계를 더 포함할 수 있고, 분류 모델을 구축하는 단계에서 트레이닝에 적용되는 입력 변수는 제1 입력 변수 세트에 대응할 수 있다.

[0006] 다른 일실시예에 따르면, 위성자료를 관측 시점이 서로 다른 제1 위성자료 및 제2 위성자료로 구분하는 단계, 제2 위성자료에서 실제 성층권 침투 대류운의 출현 및 비출현 지역을 지정하는 단계, 트레이닝을 통해 구축되는 분류 모델을 제2 위성자료에 적용하여 성층권 침투 대류운의 출현 및 비출현 지역을 분류하는 단계 및 지정하는 단계의 실제 성층권 침투 대류운의 출현 및 비출현 지역과 분류하는 단계의 성층권 침투 대류운의 출현 및 비출현 지역을 비교하여 분류 모델을 검증하는 단계를 더 포함하고 탐색하는 단계에서 탐색하는 위성자료는 제1 위성자료에 대응할 수 있다.

[0007] 또 다른 일실시예에 따르면, 탐색하는 단계는 위성자료가 위성에서 수신되는 가시영상 및 수치일기예보 대류권계면 온도자료에 해당할 수 있다.

[0008] 또 다른 일실시예에 따르면, 가시영상의 광휘값을 반사도로 변환하는 단계를 더 포함할 수 있다.

[0009] 또 다른 일실시예에 따르면, 추출하는 단계는 위성에서 수신되는 적외영상의 데이터로부터 입력 변수를 추출할 수 있다.

[0010] 또 다른 일실시예에 따르면, 적외영상의 광휘값을 밝기온도로 변환하는 단계를 더 포함할 수 있다.

[0011] 또 다른 일실시예에 따르면, 학습기법은 DT(Decision Trees), RF(Random Forest), ERT(Extremely Randomized Trees), 및 SVM(Support Vector Machines)을 포함할 수 있다.

[0012] 또 다른 일실시예에 따르면, 테스트는 통계치를 이용한 성능 평가를 통해 수행될 수 있다.

[0013] 또 다른 일실시예에 따르면, 추출하는 단계는 입력 변수를 윈도 사이즈 별로 추출할 수 있다.

[0014] 다른 일측에 따르면, 위성자료에서 성층권 침투 대류운의 출현 및 비출현의 기준이 되는 지역을 탐색하는 탐색부, 기준이 되는 지역에 대한 정보로부터 성층권 침투 대류운의 출현 및 비출현을 판단하기 위한 입력 변수를 추출하는 변수 추출부 및 입력 변수를 적용하여 학습기법을 이용한 트레이닝을 통해 성층권 침투 대류운의 출현 및 비출현의 지역을 분류하는 분류 모델을 구축하는 모델 생성부를 포함하는 분류 모델 구축 장치가 제공될 수 있다.

[0015] 일실시예에 따르면, 입력 변수를 제1 입력 변수 세트 및 제2 입력 변수 세트로 임의로 분류하는 변수 분류부 및 제2 입력 변수 세트를 적용하여 트레이닝을 통해 구축되는 분류 모델을 테스트하는 테스트부를 더 포함할 수 있

고, 모델 생성부가 트레이닝에 적용하는 입력 변수는 제1 입력 변수 세트에 대응하는 분류 모델 구축 장치가 제공될 수 있다.

[0016] 다른 일실시예에 따르면, 위성자료를 관측 시점이 서로 다른 제1 위성자료 및 제2 위성자료로 구분하여 분류하는 위성자료 분류부, 제2 위성자료에서 실제 성층권 침투 대류운의 출현 및 비출현 지역을 지정하는 지정부, 트레이닝을 통해 구축되는 분류 모델을 제2 위성자료에 적용하여 성층권 침투 대류운의 출현 및 비출현 지역을 분류하는 지역 분류부 및 지정부의 실제 성층권 침투 대류운의 출현 및 비출현 지역과 지역 분류부의 성층권 침투 대류운의 출현 및 비출현 지역을 비교하여 분류 모델을 검증하는 검증부를 더 포함할 수 있고, 탐색부가 탐색하는 위성자료는 제1 위성자료에 대응할 수 있다.

[0017] 또 다른 일실시예에 따르면, 탐색부는 위성자료가 위성에서 수신되는 가시영상 및 수치일기예보 대류권계면 온도자료에 해당할 수 있다.

[0018] 또 다른 일실시예에 따르면, 가시영상의 광휘값을 반사도로 변환하는 제1 변환부를 더 포함할 수 있다.

[0019] 또 다른 일실시예에 따르면, 변수 추출부는 위성에서 수신되는 적외영상의 데이터로부터 입력 변수를 추출할 수 있다.

[0020] 또 다른 일실시예에 따르면, 적외영상의 광휘값을 밝기온도로 변환하는 제2 변환부를 더 포함할 수 있다.

[0021] 또 다른 일실시예에 따르면, 학습기법은 DT(Decision Trees), RF(Random Forest), ERT(Extremely Randomized Trees) 및 SVM(Support Vector Machines)을 포함할 수 있다.

[0022] 또 다른 일실시예에 따르면, 테스트는 통계치를 이용한 성능 평가를 통해 수행될 수 있다.

[0023] 또 다른 일실시예에 따르면, 변수 추출부는 입력 변수를 윈도 사이즈 별로 추출할 수 있다.

[0024] 또 다른 일측에 따르면, 분류 모델 구축 방법을 수행하는 프로그램을 수록한 컴퓨터 판독 가능 기록 매체가 제공될 수 있다.

도면의 간단한 설명

[0025] 도 1은 일실시예에 따른 성층권 침투 대류운 탐지 과정을 설명하기 위한 전체 시스템도이다.

도 2는 일실시예에 따른 성층권 침투 대류운 출현 지역의 분류 모델을 구축하는 과정을 도시하는 흐름도이다.

도 3은 일실시예에 따른 성층권 침투 대류운 출현 지역의 분류 모델을 구축하는 과정에서 출현 및 비출현의 기준이 되는 지역을 지정한 데이터를 도시한다.

도 4는 일실시예에 따른 성층권 침투 대류운 출현 지역의 분류 모델을 실행하여 출현 지역을 분류한 데이터를 도시한다.

도 5는 일실시예에 따른, 도 3 및 도 4의 데이터를 비교하여 성층권 침투 대류운 출현 지역의 분류 모델을 검증하는 과정을 도시한다.

도 6은 일실시예에 따른 성층권 침투 대류운 출현 지역의 분류 모델을 구축하는 전체 과정을 도시하는 흐름도이다.

발명을 실시하기 위한 구체적인 내용

[0026] 이하에서, 실시예들을 첨부된 도면을 참조하여 상세하게 설명한다. 그러나, 권리범위는 이러한 실시예들에 의해 제한되거나 한정되는 것은 아니다. 각 도면에 제시된 동일한 참조 부호는 동일한 부재를 나타낸다.

[0027] 아래 설명에서 사용되는 용어는, 연관되는 기술 분야에서 일반적이고 보편적인 것으로 선택되었으나, 기술의 발달 및/또는 변화, 관례, 기술자의 선호 등에 따라 다른 용어가 있을 수 있다. 따라서, 아래 설명에서 사용되는 용어는 기술적 사상을 한정하는 것으로 이해되어서는 안 되며, 실시예들을 설명하기 위한 예시적 용어로 이해되어야 한다.

[0028] 또한 특정한 경우는 출원인이 임의로 선정한 용어도 있으며, 이 경우 해당되는 설명 부분에서 상세한 그 의미를 기재할 것이다. 따라서 아래 설명에서 사용되는 용어는 단순한 용어의 명칭이 아닌 그 용어가 가지는 의미와 명세서 전반에 걸친 내용을 토대로 이해되어야 한다.

[0030] 용어의 설명

[0031] 성층권 침투 대류운(Overshooting Top)이란 적란운 내의 상승기류가 아주 강할 때, 적란운의 최상층이 대류권계면을 뚫고 성층권까지 진입한 것을 말한다. 성층권 침투 대류운은 그 지역에서 강한 상승력을 지닌 대기로 인해 성층권 하부로 수증기가 침투하면서 발생하게 되고, 성층권 침투 대류운이 발생하는 주변에는 악기상 조건이 나타날 확률이 높다. 따라서 성층권 침투 대류운의 발생 지역을 탐지하는 것은 항공 기상 정보에 유용하게 이용될 수 있다.

[0032] 기계학습(Machine Learning)이란 인공지능의 한 분야로, 정보의 학습 작업을 반복적으로 수행함으로써, 빅데이터를 분석해 이후를 예측하는 응용이 가능하다. 기계학습의 예로, DT(Decision Trees), RF(Random Forest), ERT(Extremely Randomized Trees)등이 있다.

[0033] 판단 트리(Decision Trees)는 시스템을 계획 및 추진해갈 때 일종의 결정이 이루어지는데, 그 결정을 추진하면서 다양한 방안을 도출할 수 있고, 그 중 하나의 방안을 결정하면서 불확정적인 문제가 발생할 수 있으며, 결정한 방안에 의한 결과의 예상까지 고려해야 하는데 이러한 사항들을 나무 가지와 같이 도시하는 과정을 나타낸다.

[0034] 랜덤 포레스트(Random Forest)는 판단 트리의 단점을 개선한 알고리즘으로, 다수의 판단 트리를 결합하여 하나의 모형을 생성하는 방법으로 임의성(Randomness)을 관측치(Observation, instance) 뿐만 아니라 변수(feature, variable)에도 적용한 것이다. 판단 트리 표본의 개수를 조정하여 다양성을 확보하는 것이 가능하다.

[0035] ERT(Extremely Randomized Trees) 역시 새로운 트리 기반의 앙상블 방법으로, 트리 노드를 분할하는 선택이 매우 무작위적으로 구성되게 된다. 학습 샘플 변수의 출력 값과 독립되는 완전히 무작위한 트리를 만들 수 있다는 강점이 있다. 따라서 보다 정확할 수 있고, 계산의 효율성이 높다고 할 수 있다.

[0036] SVM(Support Vector Machines)은 분류(Classification) 기술의 한 종류로서, 통계 학습 이론(Statistical Learning Theory)에 근간을 두고 있다. 예시로 공간을 분류하는 데에 있어서 최대의 마진(Margin)을 가지는 평면을 찾아가는 과정을 통해 분류하는 방법을 찾는 과정에 해당한다.

[0038] 도 1은 일실시예에 따른 성층권 침투 대류운 탐지 과정을 설명하기 위한 전체 시스템도이다.

[0039] 수신부(120)는 위성(110)으로부터 위성자료를 수신하여 성층권 침투 대류운 출현 지역의 분류 모델 구축 장치(130)로 위성자료를 전송할 수 있다.

[0040] 위성(110)은 임의의 모든 위성이 될 수 있다. 수신부(120)가 수신하는 위성자료는 가시영상, 수치일기예보 온도자료, 적외영상 등 임의의 모든 위성자료가 될 수 있다.

[0041] 분류 모델 구축 장치(130)에서 생성된 분류 모델을 임의의 위성 자료에 적용하여 성층권 침투 대류운 출현 지역의 분류 결과 데이터(140)를 얻을 수 있다.

[0042] 분류 모델 구축 장치(130)는 일실시예로 탐색부, 변수 추출부 및 모델 생성부를 포함할 수 있다. 탐색부는 수신부(120)가 수신하는 위성자료에서 성층권 침투 대류운의 출현 및 비출현의 기준이 되는 지역을 탐색할 수 있다. 변수 추출부는 기준이 되는 지역에 대한 정보로부터 성층권 침투 대류운의 출현여부를 판단하기 위한 입력 변수를 추출할 수 있다. 모델 생성부는 변수 추출부가 추출하는 입력 변수를 학습기법에 적용하여 트레이닝을 통해 성층권 침투 대류운의 출현 및 비출현 지역을 분류하는 분류 모델을 구축할 수 있다. 일례로, 적용되는 학습기법에는 DT(Decision Trees), RF(Random Forest), ERT(Extremely Randomized Trees), SVM(Support Vector Machines)등이 있을 수 있다.

[0043] 분류 모델 구축 장치(130)에 대한 다른 실시예로 탐색부, 변수 추출부, 모델 생성부, 변수 분류부, 테스트부, 위성자료 분류부, 지정부, 지역 분류부, 검증부, 제1 변환부 및 제2 변환부를 포함할 수 있다. 위성자료 분류부는 수신부(120)가 수신하는 위성자료를 관측 시점이 서로 다른 제1 위성자료 및 제2 위성자료로 구분하여 분류할 수 있다. 탐색부는 제1 위성자료에서 성층권 침투 대류운의 출현 및 비출현의 기준이 되는 지역을 탐색할 수 있다. 위성자료는 가시영상, 수치일기예보 대류권계면 온도자료 및 적외영상에 해당할 수 있으나 이에 한정되지는 않는다. 일례로, 가시영상 및 수치일기예보 대류권계면 온도자료로부터 성층권 침투 대류운의 출현 및

비출현의 기준이 되는 지역을 탐색하고, 변수 추출부는 적외영상의 데이터로부터 해당 기준이 되는 지역의 데이터를 입력 변수로 추출할 수 있다. 변수 추출부는 입력 변수를 윈도 사이즈(Window size) 별로 추출하여 성층권 침투 대류운의 다양한 크기를 반영하여 입력 변수를 추출할 수 있다. 제1 변환부는 가시영상의 광휘값(Radiance)을 반사도(Reflectance)로 변환할 수 있다. 제2 변환부는 적외영상의 광휘값(Radiance)을 밝기온도로 변환할 수 있다. 탐색부 및 변수 추출부는 제1 변환부 및 제2 변환부로부터 변환된 위성자료를 이용할 수 있다. 추출되는 입력 변수를 변수 분류부는 제1 입력 변수 세트 및 제2 입력 변수 세트로 분류할 수 있다. 모델 생성부는 변수 추출부가 추출하는 제1 입력 변수 세트를 학습기법에 적용하여 트레이닝을 통해 성층권 침투 대류운의 출현 및 비출현 지역을 분류하는 분류 모델을 구축할 수 있다. 테스트부는 제2 입력 변수 세트를 적용하여 트레이닝을 통해 구축되는 분류 모델을 테스트할 수 있다. 테스트는 통계치를 이용한 성능 평가를 통해 수행될 수 있으며 일례로, PA(Producer's Accuracy), UA(User's Accuracy), OA(Overall Accuracy), Kappa(Kappa Coefficient)등이 있을 수 있으며, 이에 국한되지 않는다.

[0044] 생성되는 분류 모델을 검증하기 위하여 지정부는 제2 위성자료에서 실제 성층권 침투 대류운의 출현 및 비출현 지역을 지정할 수 있다. 이를 일례로 도 3에 도시하였다. 지역 분류부는 트레이닝을 통해 구축되는 분류 모델을 제2 위성자료에 적용하여 성층권 침투 대류운의 출현 및 비출현 지역을 분류한 결과 데이터를 얻을 수 있다. 결과 데이터는 분류 맵의 형태로 나타날 수 있으며 이는 도 4에 일례로 도시하였다. 검증부는 지정부로부터 지정되는 성층권 침투 대류운의 실제 출현 및 비출현 지역과 지역 분류부로부터 얻어지는 분류 모델의 적용에 따른 성층권 침투 대류운의 출현 및 비출현 지역의 분류 결과를 비교하여 분류 모델의 성능을 검증할 수 있다. 지정부로부터 얻어지는 실제 출현 지역에 대한 결과 데이터와 지역 분류부로부터 얻어지는 분류 모델의 실행에 따른 출현 지역에 대한 결과 데이터가 일치할수록 성층권 침투 대류운의 출현 및 비출현 지역에 대한 분류 모델의 성능은 높다고 할 수 있다.

[0045] 도 2는 일실시예에 따른 성층권 침투 대류운 출현 지역의 분류 모델을 구축하는 과정을 도시하는 흐름도이다.

[0046] 성층권 침투 대류운 출현 지역의 분류 모델을 구축하기 위해 데이터 베이스(210)를 훈련 즉, 트레이닝에 이용되는 데이터와 구축된 모델을 테스트하기 위한 데이터로 분류하는 과정(220)을 진행한다. 데이터 베이스(210)는 성층권 침투 대류운을 탐지하기 위해 얻을 수 있는 다양한 위성자료가 될 수 있다.

[0047] 훈련 및 테스트에 구분하여 적용하기 위해 데이터를 제1 입력 변수 세트 및 제2 입력 변수 세트로 분류할 수 있다. 제1 입력 변수 세트는 기계학습을 통해 성층권 침투 대류운 출현 지역의 분류 모델을 구축하기 위해 기계학습 방법을 적용한 훈련에 이용된다. 성층권 침투 대류운 출현 지역에 대한 제1 입력 변수들을 기계학습 기법을 통해 반복적으로 학습하여 분류 모델을 구축할 수 있다. 제2 입력 변수 세트는 제1 입력 변수들을 이용하여 구축되는 분류 모델을 테스트하기 위해 이용될 수 있다.

[0048] 다음으로, 도 3 내지 도 5를 통해 분류 모델을 검증하는 과정을 상세히 설명하도록 하겠다.

[0049] 도 3은 일실시예에 따른 성층권 침투 대류운 출현 지역의 분류 모델을 검증하는 과정에서 출현 및 비출현의 기준이 되는 지역을 지정한 데이터를 도시한다.

[0050] 도 3의 데이터(310)는 위성자료에 해당하며, 일례로 적외영상에 해당할 수 있다. 분류 모델을 구축하는 과정에서 사용되는 위성자료들의 시점과 다른 시점의 위성자료를 이용하여 분류 모델의 성능을 검증할 수 있다. 이 다른 시점의 위성 자료에서 성층권 침투 대류운의 출현 및 비출현 지역을 지정하여 나타낸다. 일례로 도 3에서와 같이 사각형의 형태로 성층권 침투 대류운의 출현 지역(320)을 나타낼 수 있다.

[0051] 도 4는 일실시예에 따른 성층권 침투 대류운 출현 지역의 분류 모델을 실행하여 출현 지역을 분류한 데이터를 도시한다.

[0052] 도 4의 데이터(410)는 도 3의 데이터(310)와 같은 시점의 위성자료에 해당할 수 있다. 이 위성자료에 대해, 구축되는 성층권 침투 대류운 출현 지역의 분류 모델을 적용하여, 분류 모델의 실행 결과 분류되는 성층권 침투 대류운 출현 지역(420)을 얻을 수 있다. 출현 지역(420)은 다양한 형태로 표현될 수 있다.

[0053] 도 5는 일실시예에 따른, 도 3 및 도 4의 데이터를 비교하여 성층권 침투 대류운 출현 지역의 분류 모델을 검증하는 과정을 도시한다.

[0054] 도 5의 데이터(510) 도 3 및 도 4의 데이터(310, 410)를 종합하여 얻을 수 있다. 지정한 성층권 침투 대류운 출현 지역(520)과 분류 모델 실행 결과 분류된 성층권 침투 대류운 출현 지역(530)을 비교하여, 성층권 침투 대류운 분류 모델의 성능을 검증할 수 있다. 지정한 성층권 침투 대류운 출현 지역(520)과 분류 모델 실행 결과

분류된 성층권 침투 대류운 출현 지역(530)의 일치도가 높을수록, 분류 모델의 성능이 높다고 할 수 있다.

[0055] 도 6은 일실시예에 따른 성층권 침투 대류운 출현 지역의 분류 모델을 구축하는 전체 과정을 도시하는 흐름도이다.

[0056] 도 1의 위성(110)으로부터 수신되는 임의의 위성자료들 중 일실시예에 따라 가시영상(610), 수치일기예보 대류권계면 온도자료(630) 및 적외영상(650)의 데이터를 이용할 수 있다. 가시영상(610)은 오리지널 값이 광휘값(Radiance)으로 주어지므로 이를 반사도(Reflectance)로 변환하는 과정(615)이 필요할 수 있다. 변환된 가시영상의 데이터와 수치일기예보 대류권계면 온도자료로부터 성층권 침투 대류운 출현(OT) 및 비출현(Non-OT)의 기준 지역을 탐색할 수 있다.

[0057] 적외영상(650)은 오리지널 값이 광휘값으로 주어지므로 이를 밝기온도로 변환하는 과정(655)이 필요할 수 있다.

[0058] 적외영상으로부터 탐색되는 성층권 침투 대류운 출현 및 비출현의 기준 지역에 대한 입력변수를 추출(660)할 수 있다. 입력변수는 기계학습 훈련과정에서 성층권 침투 대류운의 분류 모델을 구축하기 위해 이용되는 변수로서 적외영상으로부터 추출되는 값들로 구성된 변수일 수 있다. 또한 입력변수를 윈도 사이즈(Window size) 별로 추출하여, 다양한 성층권 침투 대류운의 크기를 반영할 수 있다.

[0059] 추출되는 입력변수들 중 일부를 임의로 제1 입력 변수 세트 및 제2 입력 변수 세트로 분류하여, 제1 입력 변수 세트는 성층권 침투 대류운 출현 지역의 분류 모델을 구축하기 위한 훈련 즉 트레이닝(670)에 적용할 수 있고, 제2 입력 변수 세트는 제1 입력 변수 세트를 적용하여 구축된 성층권 침투 대류운 출현 지역의 분류 모델을 테스트(670)하는 데에 적용될 수 있다. 분류 모델을 구축하는 트레이닝은 기계학습 기법으로 이루어질 수 있다. 기계학습 기법은 특정 기법으로 한정되지 않으나 일실시예에 따라 DT(Decision Trees), RF(Random Forest), ERT(Extremely Randomized Trees), 및 SVM(Support Vector Machines)등 일 수 있다.

[0060] 추출되는 입력변수들 중 훈련 및 테스트에 적용되는 시점과 다른 시점으로부터 추출되는 입력변수들을 모델 검증용 데이터(680)에 적용할 수 있다. 모델 검증용 데이터(680)를 성층권 침투 대류운 출현 지역의 분류 모델 구축 과정(675)에서 구축된 분류 모델에 적용하여 성층권 침투 대류운 출현 지역의 분류 맵을 추출(685)할 수 있다. 추출된 분류 맵(도 3 내지 도 5의 데이터에 해당할 수 있다.)을 분석하여 구축되는 분류 모델의 성능을 검증할 수 있다. 구축되는 분류 모델을 검증(690)하는 과정은 위 도3 내지 도5에 대한 설명에서 상세히 서술하였다.

[0061] 아래에서 일실시예로 본원의 발명을 보다 상세히 설명하도록 하겠다. 아래에서 설명하는 예시나 수치, 자료 등은 어디까지나 발명의 이해를 돕기 위한 예시일 뿐이므로, 발명은 이에 국한되지 않는다.

[0063] 실시예

[0064] 반사도로 변환되는 가시영상과 수치일기예보 대류권계면 온도자료 및 밝기 온도로 변환되는 적외영상의 위성자료를 이용할 수 있다. 가시영상과 수치일기예보 대류권계면 온도자료로부터 성층권 침투 대류운 출현 지역의 기준이 되는 지역을 탐색할 수 있다. 적외영상에서 탐색되는 출현 지역의 기준 지점에 대한 입력 변수를 추출할 수 있다. 입력변수는 적외영상으로부터 얻을 수 있는 데이터로 구성될 수 있다. 일례로 입력변수는 위성의 채널로부터 확보된 주파수 채널의 정보와 위성자료의 픽셀의 중심으로부터의 차이 등에 해당할 수 있다. 아래의 표 1은 Himawari-8 위성의 AHI images 로부터 16개의 입력 변수를 추출한 예시를 나타낸다.

표 1

[0065]

Satellite/ Sensor	List of used variables (a total of 16 input variables)	Abbreviations	Period	Spatial resolution	Temporal resolution

Himawari-8/ AHI	T_B (IR 11.2 ㎛)	Tb_{11}	August 2015	2 km	10 minutes
	Mean of Tb_{11} (3 × 3 MWS)	Mean3MWS			
	Mean of Tb_{11} (5 × 5 MWS)	Mean5MWS			
	Mean of Tb_{11} (7 × 7 MWS)	Mean7MWS			
	Mean of Tb_{11} (9 × 9 MWS)	Mean9MWS			
	Mean of Tb_{11} (11 × 11 MWS)	Mean11MWS			
	STD of Tb_{11} (3 × 3 MWS)	STD3MWS			
	STD of Tb_{11} (5 × 5 MWS)	STD5MWS			
	STD of Tb_{11} (7 × 7 MWS)	STD7MWS			
	STD of Tb_{11} (9 × 9 MWS)	STD9MWS			
	STD of Tb_{11} (11 × 11 MWS)	STD11MWS			
	Difference between the center of 3 × 3 MWS and its boundary pixels	Diff3MWS			
	Difference between the center of 5 × 5 MWS and its boundary pixels	Diff5MWS			
	Difference between the center of 7 × 7 MWS and its boundary pixels	Diff7MWS			
	Difference between the center of 9 × 9 MWS and its boundary pixels	Diff9MWS			
	Difference between the center of 11 × 11 MWS and its boundary pixels	Diff11MWS			

[0066] 일례를 나타낸 것으로, 본원의 입력 변수의 개수는 16개로 한정되지 않는다. 추출되는 입력 변수의 데이터를 임의로 80:20의 양으로 분류하여, 각각 모델을 구축하기 위한 훈련 자료(위 설명에서 제1 입력 변수 세트에 해당한다.) 및 구축되는 모델을 평가하기 위한 테스트 자료(위 설명에서 제2 입력 변수 세트에 해당한다.)에 이용할 수 있다.

[0067] 훈련 자료 및 테스트 자료를 입력 변수로 하여, 기계학습을 통해 분류 모델을 구축할 수 있다. 예시로, 기계학습 기법은 DT(Decision Trees), RF(Random Forest), ERT(Extremely Randomized Trees), SVM(Support Vector Machines) 등이 해당할 수 있다.

[0068] 구축되는 분류 모델들에 대한 테스트는 통계치를 이용할 수 있다. 통계치의 예시로는 PA(Producer's Accuracy), UA(User's Accuracy), OA(Overall Accuracy), Kappa(Kappa Coefficient)등이 있을 수 있으며, 이에 국한되지 않는다. 아래 표 2 는 Decision Trees 기법을 통해 구축되는 분류 모델을 통계치를 이용해 성능 평가한 결과를 나타낸다.

표 2

[0069]

Reference / Classified as	OT	non-OT	Sum	User's Accuracy
OT	82	8	90	91.11%
non-OT	2	255	257	99.22%
Sum	84	263	347	
Producer's accuracy	97.62%	96.96%		
Overall accuracy	97.12%			
Kappa coefficient	0.923			

[0070] 아래 표 3 는 Random Forest 기법을 통해 구축되는 분류 모델을 통계치를 이용해 성능 평가한 결과를 나타낸다.

표 3

[0071]

Reference / Classified as	OT	non-OT	Sum	User's Accuracy
OT	81	4	85	95.29%
non-OT	3	259	262	98.85%
Sum	84	263	347	
Producer's accuracy	96.43%	98.48%		
Overall accuracy	97.98%			
Kappa coefficient	0.945			

[0072] 아래 표 4 는 Extremely Randomized Trees 기법을 통해 구축되는 분류 모델을 통계치를 이용해 성능 평가한 결과를 나타낸다.

표 4

[0073]

Reference / Classified as	OT	non-OT	Sum	User's Accuracy
OT	81	3	84	96.43%
non-OT	3	260	263	98.86%
Sum	84	263	347	
Producer's accuracy	96.43%	98.86%		
Overall accuracy	98.27%			
Kappa coefficient	0.953			

[0074] 아래 표 5 는 Support Vector Machines 기법을 통해 구축되는 분류 모델을 통계치를 이용해 성능 평가한 결과를 나타낸다.

표 5

[0075]

Reference / Classified as	OT	non-OT	Sum	User's Accuracy
OT	80	0	80	100%
non-OT	4	263	267	98.50%
Sum	84	263	347	
Producer's accuracy	95.24%	100%		
Overall accuracy	98.85%			
Kappa coefficient	0.968			

[0076] 기계학습 기법 중 Decision Trees 를 통해 생성되는 분류 모델은 if-then 룰 형식으로 분류 모델이 만들어지게 되어, 이보다 트리 개수가 훨씬 많은 Random Forest 혹은 Extremely Randomized Trees 나, 기계학습기법이 아닌 Support Vector Machines 보다 쉽게 이해할 수 있는 모델로 구축되게 될 수 있다. 아래 표 6은 Decision Trees 를 통해 구축되는 분류 모델에서 분류 기준을 나타내며, 왼쪽의 룰은 성층권 침투 대류운 출현 지역의 분류 룰을 나타내고, 오른쪽의 룰은 성층권 침투 대류운 비출현 지역의 분류 룰을 나타낸다.

표 6

[0077]

Classified as / Rule	OT	non-OT
1	STD (3 MWS) > 0.934474 Mean (5 MWS) ≤≤ 193.0878	Mean (5 MWS) > 193.0878 STD (7 MWS) ≤≤ 1.70502

2	Mean (3 MWS) ≤≤ 191.1089	Mean (3 MWS) > 191.1089 STD (3 MWS) ≤≤ 0.934474
3	Tb_{11} ≤≤ 196.0474 STD (3 MWS) > 0.934474 STD (7 MWS) > 1.70502	Tb_{11} > 196.0474

[0078] 구축되는 분류 모델은 훈련이나 테스트 자료에 사용되지 않은 다른 시점의 위성자료를 통해 검증되는 과정을 거칠 수 있다. 검증되는 과정은 도 3 내지 도 5를 통해 상세히 설명하였다.

[0079] 이상에서 설명된 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPA(field programmable array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

[0080] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장 매체 또는 장치, 또는 전송되는 신호 파(signal wave)에 영구적으로, 또는 일시적으로 구체화(embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

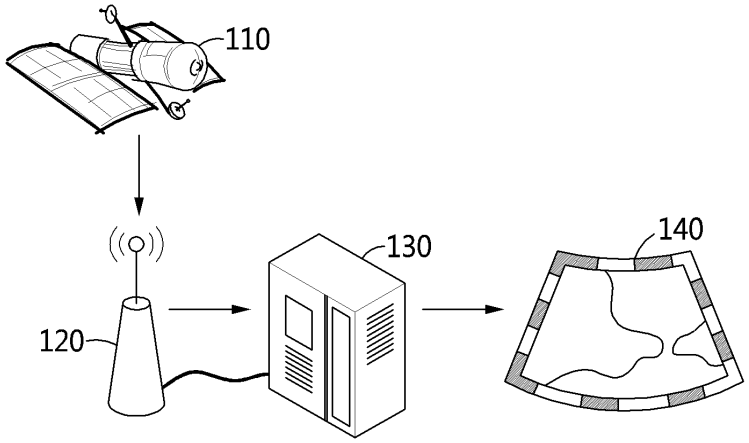
[0081] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다. 상기된 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

[0082] 실시예들이 비록 한정된 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.

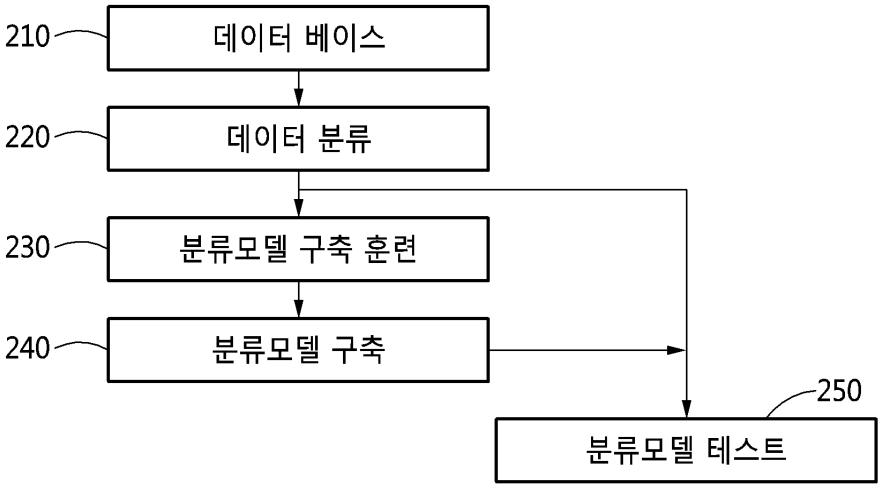
[0083] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

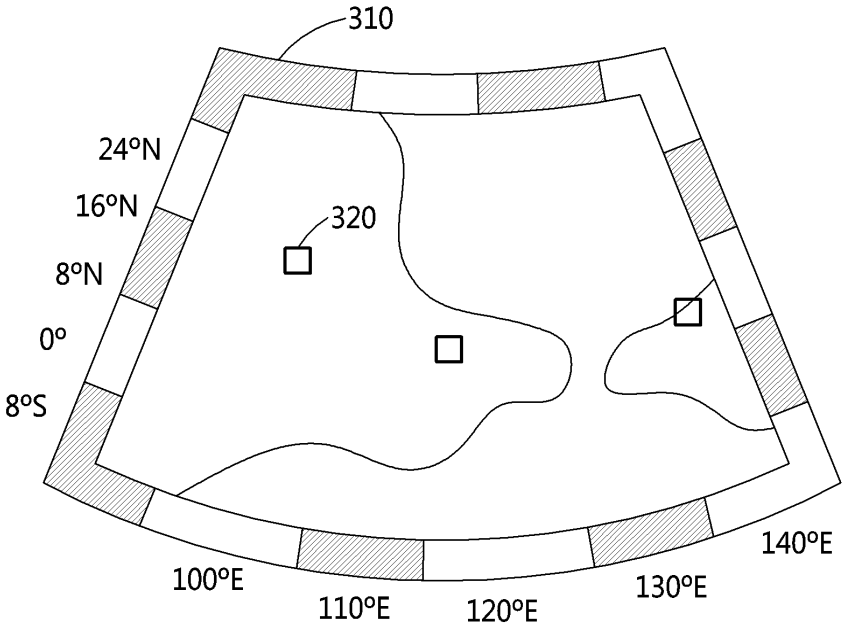
도면1


110
120
130
140

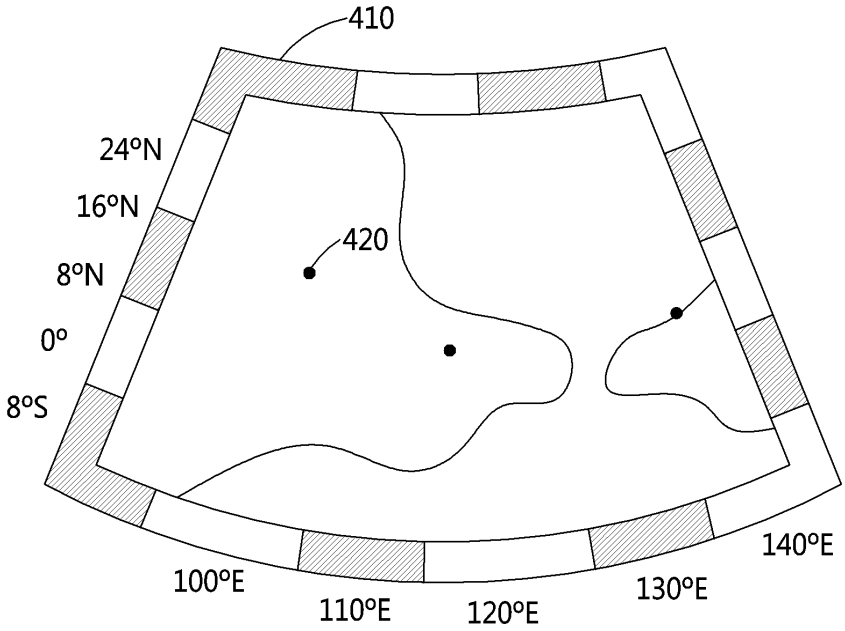
도면2


210
데이터 베이스
220
데이터 분류
230
분류모델 구축 훈련
240
분류모델 구축
250
분류모델 테스트

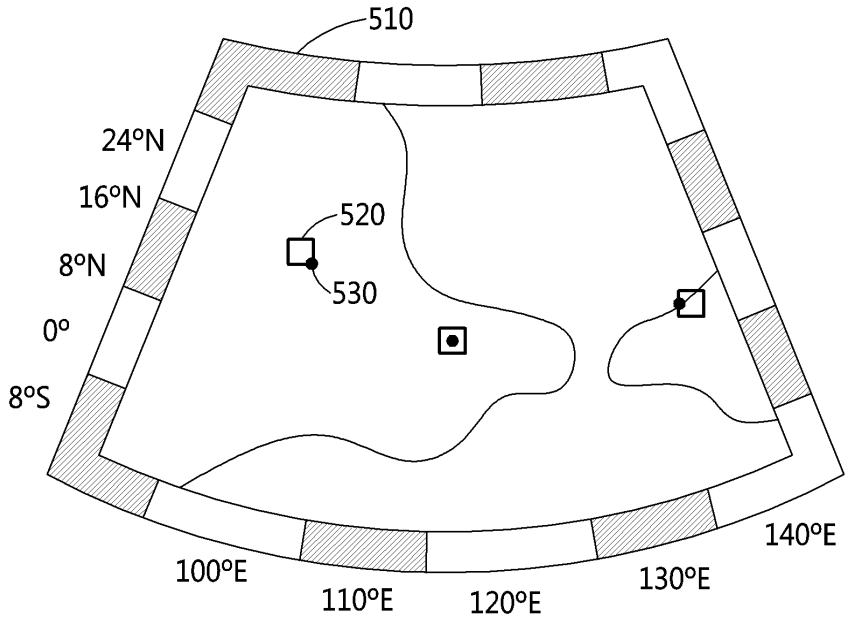
도면3


310
24°N
16°N
8°N
0°
8°S
320
100°E
110°E
120°E
130°E
140°E

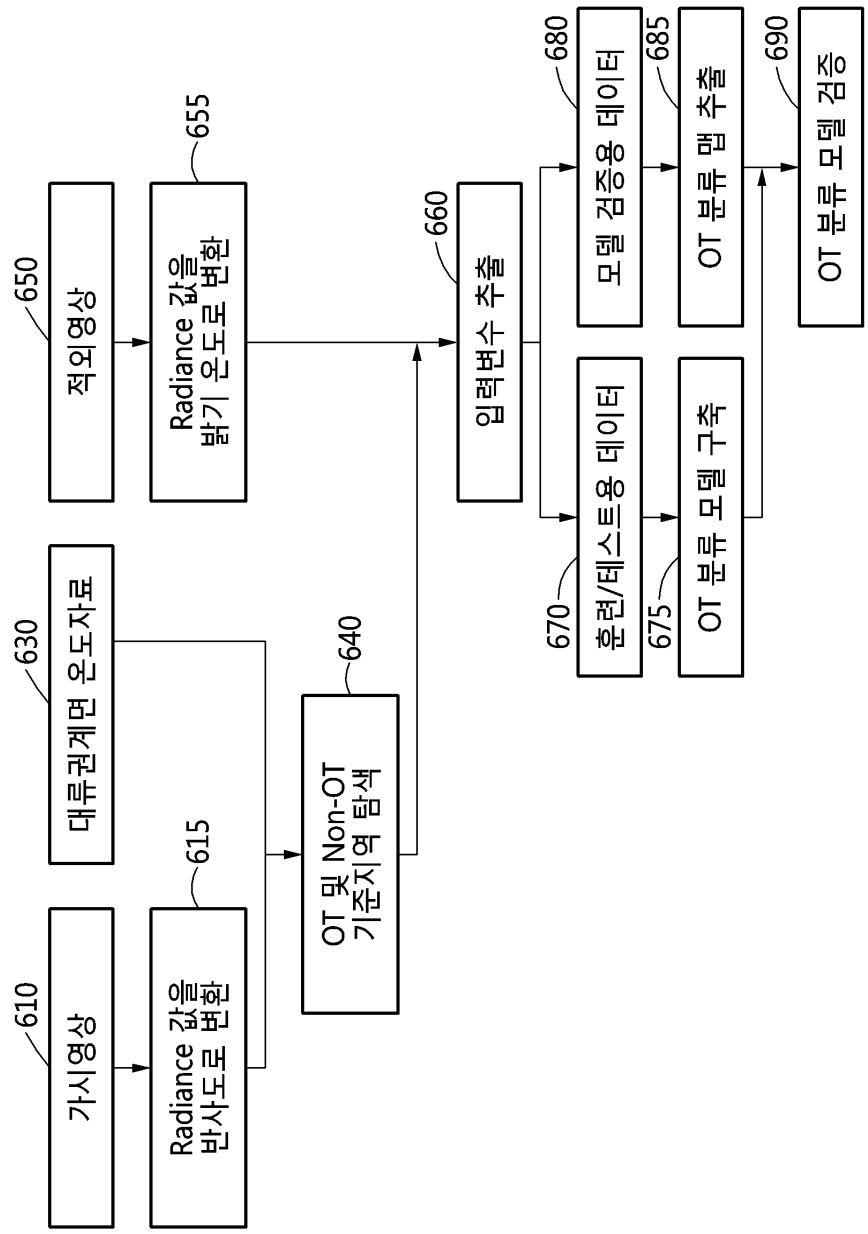
도면4


410
24°N
16°N
8°N
0°
8°S
420
100°E
110°E
120°E
130°E
140°E

도면5



도면6


610
가시영상
615
Radiance 값을 반사도로 변환
630
대류권계면 온도자료
640
OT 및 Non-OT 기준지역 탐색
650
적외영상
655
Radiance 값을 밝기 온도로 변환
660
입력변수 추출
670
훈련/테스트용 데이터
675
OT 분류 모델 구축
680
모델 검증용 데이터
685
OT 분류 맵 추출
690
OT 분류 모델 검증