(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2025년04월14일
(11) 등록번호 10-2795932
(24) 등록일자 2025년04월10일

(51) 국제특허분류(Int. Cl.)
G06N 20/20 (2019.01) *G01N 33/24* (2006.01)
G06N 3/0464 (2023.01)
(52) CPC특허분류
G06N 20/20 (2021.08)
G01N 33/24 (2019.01)
(21) 출원번호 10-2024-0082178
(22) 출원일자 2024년06월24일
심사청구일자 2024년06월24일
(56) 선행기술조사문헌
"Rapid and Accurate Prediction of Soil Texture Using an Image-Based Deep Learning Autoencoder Convolutional Neural Network Random Forest Algorithm", Agronomy, 2022.12.03.*
US20240087311 A1*
CN117152733 A
US20200383284 A1
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
세종대학교산학협력단
서울특별시 광진구 능동로 209 (군자동, 세종대학교)
(72) 발명자
최수미
서울특별시 광진구 능동로 209, 대양AI센터 720호(군자동)
아볼가셈
서울특별시 광진구 능동로 209, 대양AI센터 433호(군자동)
(뒷면에 계속)
(74) 대리인
양성보

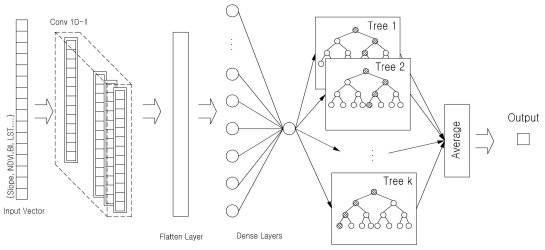
전체 청구항 수 : 총 7 항 심사관 : 박승철

(54) 발명의 명칭 **토양의 물리적 특성 예측을 위한 지리공간 인공지능과 위성영상의 융합 방법 및 장치**

(57) 요 약

토양의 물리적 특성 예측을 위한 지리공간 인공지능과 위성영상의 융합 방법 및 장치를 제공한다. DL(deep learning)과 ML(machine learning)의 하이브리드 네트워크로서 CNN(Convolutional Neural Network) 알고리즘과 RF(Random Forest) 알고리즘이 결합된 하이브리드 CNN-RF 모델을 통해 토양의 물리적 특성을 예측할 수 있다.

대 표 도



(52) CPC특허분류
G06N 3/0464 (2023.01)
(72) 발명자
서명배
서울특별시 광진구 능동로 209, 광개토관 1001호(군자동)
바히드 라자비 테르메
서울특별시 광진구 능동로 209, 광개토관 1001호(군자동)
김용국
서울특별시 광진구 능동로 209, 대양AI센터 722호(군자동)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711179366
과제번호	RS-2022-00156354
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	대학ICT연구센터지원사업
연구과제명	실-가상 연계 메타버스를 위한 초실감 XR 기술 연구
과제수행기관명	세종대학교 산학협력단
연구기간	2022.07.01 ~ 2029.12.31

공지예외적용 : 있음

명 세 서

청구범위

청구항 1

적어도 하나의 프로세서를 포함하는 컴퓨터 장치의 토양 특성 예측 방법에 있어서,

상기 적어도 하나의 프로세서에 의해, DL(deep learning)과 ML(machine learning)의 하이브리드 네트워크로서 CNN(Convolutional Neural Network) 알고리즘과 RF(Random Forest) 알고리즘이 결합된 하이브리드 CNN-RF 모델을 통해 토양의 물리적 특성을 예측하는 단계

를 포함하고,

상기 하이브리드 CNN-RF 모델은 RS(Remote sensing) 매개변수와 지형 매개변수(Topographic Parameter) 및 기후 매개변수(Climatic Parameter)를 포함하고,

상기 RS 매개변수는 위성 영상에서 추출되는 데이터로, Landsat-8의 밴드 1(B1) 내지 밴드 5(B5), 및 밴드 7(B7), BI(Brightness Index), CI(Coloration Index), CLI(Clay Index), EVI(Enhanced Vegetation Index), LST(Land Surface Temperature), HI(Hue Index), NDVI(Normalized Difference Vegetation Index), RI(Redness Index), 및 SI(Saturation Index)가 포함되고,

상기 지형 매개변수는 지리정보시스템(GIS)의 지형 모델에서 추출되는 데이터로, 측면(aspect), 고도(elevation), 경사(slope), DR(duration radiation), MRVBF(multi-resolution index of valley bottom flatness), MRRTF(multi-resolution ridgetop flatness index), 및 TWI(topographic wetness index)가 포함되고,

상기 하이브리드 CNN-RF 모델은 mXn(m은 토양 샘플의 양, n은 토양 질감 속성에 영향을 미치는 매개변수의 수) 구조의 입력 행렬을 포함하고,

상기 예측하는 단계는,

상기 하이브리드 CNN-RF 모델을 사용하여 대상 지역의 토양 질감 특성에 대한 예측 지도(predictive map)를 생성하는 것으로,

CNN 모델의 히든 레이어(hidden layer)를 통해 입력 정보에서 공간 패턴과 컨텍스트 정보를 포함한 특징을 추출하는 단계;

회색조 동시 발생 행렬(gray-level cooccurrence matrix)을 이용하여 대상 지역의 균질성, 대비, 상이성, 엔트로피를 포함한 특징을 추출하는 단계; 및

RF 알고리즘을 통해 상기 추출된 특징에 대한 회귀 분석을 수행하는 단계

를 포함하는 토양 특성 예측 방법.

청구항 2

제1항에 있어서,

상기 예측하는 단계는,

대상 지역의 토양 샘플 데이터를 이용하여 비중계법(hydrometer method)을 통해 점토(clay), 모래(sand), 실트(silt)를 포함한 토양 질감 특성을 모델링하는 단계

를 포함하는 토양 특성 예측 방법.

청구항 3

제2항에 있어서,

상기 토양 샘플 데이터는 이상치 제거를 거쳐 축소되는 것

을 특징으로 하는 토양 특성 예측 방법.

청구항 4

삭제

청구항 5

제1항에 있어서,

상기 예측하는 단계는,

위성 영상에서 추출된 매개변수를 이용하여 공간 데이터베이스를 생성하는 단계

를 포함하는 토양 특성 예측 방법.

청구항 6

삭제

청구항 7

삭제

청구항 8

삭제

청구항 9

삭제

청구항 10

삭제

청구항 11

제1항 내지 제3항, 제5항 중 어느 한 항의 토양 특성 예측 방법을 상기 컴퓨터 장치에 실행시키기 위해 컴퓨터 판독 가능한 기록매체에 저장된 컴퓨터 프로그램.

청구항 12

컴퓨터 장치에서 판독 가능한 명령을 실행하도록 구현되는 적어도 하나의 프로세서

를 포함하고,

상기 적어도 하나의 프로세서는,

DL(deep learning)과 ML(machine learning)의 하이브리드 네트워크로서 CNN(Convolutional Neural Network) 알고리즘과 RF(Random Forest) 알고리즘이 결합된 하이브리드 CNN-RF 모델을 통해 토양의 물리적 특성을 예측하고,

상기 하이브리드 CNN-RF 모델은 RS(Remote sensing) 매개변수와 지형 매개변수(Topographic Parameter) 및 기후 매개변수(Climatic Parameter)를 포함하고,

상기 RS 매개변수는 위성 영상에서 추출되는 데이터로, Landsat-8의 밴드 1(B1) 내지 밴드 5(B5), 및 밴드 7(B7), BI(Brightness Index), CI(Coloration Index), CLI(Clay Index), EVI(Enhanced Vegetation Index), LST(Land Surface Temperature), HI(Hue Index), NDVI(Normalized Difference Vegetation Index), RI(Redness Index), 및 SI(Saturation Index)가 포함되고,

상기 지형 매개변수는 지리정보시스템(GIS)의 지형 모델에서 추출되는 데이터로, 측면(aspect), 고도(elevation), 경사(slope), DR(duration radiation), MRVBF(multi-resolution index of valley bottom flatness), MRRTF(multi-resolution ridgetop flatness index), 및 TWI(topographic wetness index)가 포함되

고,

상기 하이브리드 CNN-RF 모델은 mXn(m은 토양 샘플의 양, n은 토양 질감 속성에 영향을 미치는 매개변수의 수) 구조의 입력 행렬을 포함하고,

상기 적어도 하나의 프로세서는,

상기 하이브리드 CNN-RF 모델을 사용하여 대상 지역의 토양 질감 특성에 대한 예측 지도(predictive map)를 생성하는 것으로,

CNN 모델의 히든 레이어(hidden layer)를 통해 입력 정보에서 공간 패턴과 컨텍스트 정보를 포함한 특징을 추출하고,

회색조 동시 발생 행렬(gray-level cooccurrence matrix)을 이용하여 대상 지역의 균질성, 대비, 상이성, 엔트로피를 포함한 특징을 추출하고,

RF 알고리즘을 통해 상기 추출된 특징에 대한 회귀 분석을 수행하는 것

을 특징으로 하는 컴퓨터 장치.

청구항 13

삭제

청구항 14

제12항에 있어서,

상기 적어도 하나의 프로세서는,

위성 영상에서 추출된 매개변수를 이용하여 토양 질감 특성을 모델링하기 위한 공간 데이터베이스를 생성하는 것

을 특징으로 하는 컴퓨터 장치.

청구항 15

삭제

발명의 설명

기 술 분 야

[0001] 아래의 설명은 토양의 물리적 특성을 예측하는 기술에 관한 것이다.

배 경 기 술

[0002] 토양은 기후와 생태계 조절의 중요한 구성 요소이자 인류 식량의 97%를 생산하는 기본 요소이다. 토양은 또한 농업 생산성, 유역 보호, 환경 및 야생동물에 큰 영향을 미친다. 토양 질감은 토양 침식, 물 이동, 품질 관리 및 생산성에 매우 중요하다. 토양 질감의 입자 크기 분류에는 모래(2~0.05㎜), 미사(0.05~0.002㎜), 점토(<0.002㎜)가 포함된다. 토양이 직면한 중요한 과제 중에는 다양한 규모의 토양 침식과 강우 침식이 있으며, 이는 토양 특성, 특히 질감을 변화시킬 수 있다. 따라서 토양 특성에 대한 공간적 예측은 인간의 이용이 영향을 미치는 토양의 질을 평가하는 데 매우 중요하다.

[0003] 원격탐사(RS: remote sensing) 데이터는 전 세계적으로 이용 가능하고 농업에서 매우 가치 있는 풍부한 정보 소스이다. RS 기술의 발전으로 대규모 공간 및 시간 규모에서 데이터 처리가 크게 향상되었다. 이전에는 항공 이미지와 디지털 이미지 처리를 사용하여 농지를 모니터링했다. 그러나, RS는 추정의 정확성과 효율성을 향상시키면서 수집된 현장 데이터를 줄일 수 있다.

[0004] RS는 지리정보시스템(GIS)과 결합하여 비용, 시간, 인적 자원 측면에서 수집, 저장, 분석 및 모델링의 효율성을 높일 수 있다. 또한, GIS는 공간 예측을 돕기 위해 공간 정보와 환경 매개변수를 결합하는 다양한 도구를 제공한다. GIS는 매핑, 데이터 관리, 공간 분석을 위한 효과적인 분석 도구이기도 하다.

[0005] RS 및 GIS 데이터는 현상의 공간 모델링을 위한 예측 변수로 사용될 수 있다. 최근 몇 년 동안 토양 특성을 예측하는 데 공간 정보 시스템과 RS 도구 또는 기능을 사용하는 데 상당한 발전이 있었다. 토양 조직의 공간적 분포를 예측하기 위해 크리깅(kriging), 다중 단계적 회귀분석(multiple stepwise regression), 부분 최소 제곱 회귀분석(partial least squares regression), 코크리깅(cokriging)과 같은 다양한 통계 및 지리통계학적 방법이 이전에 사용되어 왔다.

[0006] 그러나, 이러한 방법은 통계적 가정에 크게 의존하고 데이터 크기가 증가함에 따라 계산 집약적이다. 이러한 한계를 극복하기 위해 토양 질감 특성을 예측하기 위해 기계 학습(ML) 알고리즘이 적용되었다. 회귀 트리(regression trees), BRT(Boosted Regression Trees), RF(Random Forest) 및 SVM(Support Vector Machine)과 같은 ML 알고리즘은 토양 질감 속성을 매핑하는 기능을 보여주었다.

[0007] ML 알고리즘은 암시적 관계를 발견하고 식별함으로써 고차원 및 다변수 데이터를 관리하는 데 상당한 이점을 제공한다. ML 알고리즘은 이러한 장점에도 불구하고 국소적으로 최적의 해만 제공하는 문제, 훈련 시간이 길어지면 성능이 저하되는 문제, 최적의 학습률을 찾는 데 어려움을 겪는 문제 등이 발생하기 쉽다.

[0008] 일부 ML 알고리즘은 데이터 양이 증가함에 따라 성능 포화 상태를 나타낼 수 있지만, 데이터 양과 알고리즘 성능 간의 관계는 데이터의 이질성 및 관련성을 포함한 다양한 요인의 영향을 받는다. 데이터가 다양하고 다양한 규모와 맥락에 걸쳐 중요한 정보를 포함하는 경우 데이터 양을 늘리면 모델 성능이 향상될 수 있다. 그러나, 추가 볼륨이 모델 훈련 및 일반화에 의미 있게 기여하도록 데이터를 신중하게 선별하고 사전 처리하는 것이 중요하다. 더욱이, ML 알고리즘은 관련이 없고 중복된 정보를 감지할 수 없어 성능에 부정적인 영향을 미친다. ML은 복잡한 데이터를 처리할 수 있지만 히든 레이어(hidden layer)가 너무 많으면 과적합 및 그래디언트 소멸과 같은 문제가 발생할 수 있다.

[0009] 강력한 예측 정확도를 갖춘 딥러닝(DL)은 공간 예측에서 ML보다 성능이 뛰어나다. 복잡한 토양 문제를 해결하기 위해 DL에 기반을 둔 CNN(Convolutional Neural Network)과 같은 정교한 알고리즘을 사용하여 정확도를 높이고 불확실성을 줄일 수 있다. 또한, DL 네트워크는 ML 모델에는 없는 자동 정보 추출 기능을 제공한다. 전반적으로 DL은 향상된 성능, 자동 기능 추출 및 향상된 확장성을 제공하여 ML의 단점을 해결한다. 다양한 연구자들이 CNN 알고리즘을 사용하여 토양 질감 및 토양 염도를 예측하고 LSTM(long short-term memory) 알고리즘을 사용하여 토양 수분을 예측하는 등 토양 과학 문제를 해결하기 위해 DL 모델을 활용해 왔다. DL 모델에는 여러 가지 장점이 있지만 계산 복잡성 및 과적합과 같은 단점도 있다.

[0010] 이러한 한계를 극복하기 위해 DL 모델과 ML 알고리즘을 결합할 것을 제안한 바 있다. 이러한 결합된 네트워크에서는 DL 모델의 계층적 특성으로 인해 원시 데이터에서 필수 특징을 자동으로 추출할 수 있으며, ML 알고리즘은 DL 모델보다 회귀 작업을 더 효율적으로 처리하여 각 모델의 단점을 해결한다.

[0011] ML 및 DL 모델과 관련된 장단점에도 불구하고 이 두 가지 접근 방식의 통합은 다양한 연구 영역에서 널리 사용되고 있다. 예를 들어, CNN과 RF 조합은 조기 지진 경보 시스템(early earthquake warning systems)과 위성 이미지를 사용한 빈곤 추정(poverty estimation using satellite imagery) 등에 적용되고 있다. 또한, ML 알고리즘은 DL 신경망과 통합되어 홍수 가능성(flood potential)을 추정하고, CNN은 SVM, RF 및 논리 회귀(logical regression)와 결합하여 산사태 취약성(landslide susceptibility)을 평가하여 결과의 성능과 정확성이 향상되었다.

선행기술문헌

비특허문헌

[0012] (비특허문헌 0001) 1. Taghizadeh-Mehrjardi, R.; Khademi, H.; Khayamim, F.; Zeraatpisheh, M.; Heung, B.; Scholten, T. A comparison of model averaging techniques to predict the spatial distribution of soil properties. Remote Sens. 2022, 14, 472.

발명의 내용

해결하려는 과제

[0013] 토양 질감 특성의 공간 예측 정확도를 높이기 위해 RF와 CNN이라는 두 가지 알고리즘을 조합하여 활용할 수 있

다. 과적합을 극복하는 것 외에도 RF 알고리즘은 공간 모델링에서 다른 ML 알고리즘에 비해 허용 가능한 정확도를 보장할 수 있다. 반면, CNN 알고리즘은 컨볼루션 레이어(convolution layer)를 통한 정보 처리를 통해 다양한 특징, 특히 공간적 특징을 자동으로 추출할 수 있다.

과제의 해결 수단

[0014] 적어도 하나의 프로세서를 포함하는 컴퓨터 장치의 토양 특성 예측 방법에 있어서, 상기 적어도 하나의 프로세서에 의해, DL(deep learning)과 ML(machine learning)의 하이브리드 네트워크로서 CNN(Convolutional Neural Network) 알고리즘과 RF(Random Forest) 알고리즘이 결합된 하이브리드 CNN-RF 모델을 통해 토양의 물리적 특성을 예측하는 단계를 포함하는 토양 특성 예측 방법을 제공한다.

[0015] 일 측면에 따르면, 상기 예측하는 단계는, 대상 지역의 토양 샘플 데이터를 이용하여 비중계법(hydrometer method)을 통해 점토(clay), 모래(sand), 실트(silt)를 포함한 토양 질감 특성을 모델링하는 단계를 포함할 수 있다.

[0016] 다른 측면에 따르면, 상기 토양 샘플 데이터는 이상치 제거를 거쳐 축소될 수 있다.

[0017] 또 다른 측면에 따르면, 상기 예측하는 단계는, 상기 하이브리드 CNN-RF 모델을 사용하여 대상 지역의 토양 질감 특성에 대한 예측 지도(predictive map)를 생성하는 단계를 포함할 수 있다.

[0018] 또 다른 측면에 따르면, 상기 예측하는 단계는, 위성 영상에서 추출된 매개변수를 이용하여 공간 데이터베이스를 생성하는 단계를 포함할 수 있다.

[0019] 또 다른 측면에 따르면, 상기 하이브리드 CNN-RF 모델은 RS(Remote sensing) 매개변수와 지형 매개변수(Topographic Parameter) 및 기후 매개변수(Climatic Parameter)를 포함할 수 있다.

[0020] 또 다른 측면에 따르면, 상기 RS 매개변수는 위성 영상에서 추출되는 데이터로, Landsat-8의 밴드 1(B1) 내지 밴드 5(B5), 및 밴드 7(B7), BI(Brightness Index), CI(Coloration Index), CLI(Clay Index), EVI(Enhanced Vegetation Index), LST(Land Surface Temperature), HI(Hue Index), NDVI(Normalized Difference Vegetation Index), RI(Redness Index), SI(Saturation Index)가 포함될 수 있다.

[0021] 또 다른 측면에 따르면, 상기 지형 매개변수는 지리정보시스템(GIS)의 지형 모델에서 추출되는 데이터로, 측면(aspect), 고도(elevation), 경사(slope), DR(duration radiation), MRVBF(multi-resolution index of valley bottom flatness), MRRTF(multi-resolution ridgetop flatness index), 및 TWI(topographic wetness index)가 포함될 수 있다.

[0022] 또 다른 측면에 따르면, 상기 하이브리드 CNN-RF 모델은 mXn(m은 토양 샘플의 양, n은 토양 질감 속성에 영향을 미치는 매개변수의 수) 구조의 입력 행렬을 포함할 수 있다.

[0023] 또 다른 측면에 따르면, 상기 예측하는 단계는, 상기 하이브리드 CNN-RF 모델을 통해 토양 질감 특성에 대한 예측 값을 획득하는 것으로, CNN 모델을 통해 입력 정보에서 공간 패턴과 컨텍스트 정보를 포함한 관련 특징을 추출한 후 RF 알고리즘을 통해 상기 추출된 특징에 대한 회귀 분석을 수행할 수 있다.

[0024] 상기 토양 특성 예측 방법을 상기 컴퓨터 장치에 실행시키기 위해 컴퓨터 판독 가능한 기록매체에 저장된 컴퓨터 프로그램을 제공한다.

[0025] 컴퓨터 장치에서 판독 가능한 명령을 실행하도록 구현되는 적어도 하나의 프로세서를 포함하고, 상기 적어도 하나의 프로세서는, DL(deep learning)과 ML(machine learning)의 하이브리드 네트워크로서 CNN(Convolutional Neural Network) 알고리즘과 RF(Random Forest) 알고리즘이 결합된 하이브리드 CNN-RF 모델을 통해 토양의 물리적 특성을 예측하는 것을 특징으로 하는 컴퓨터 장치를 제공한다.

발명의 효과

[0026] 본 발명의 실시예들에 따르면, 토양 질감의 공간 예측을 위한 RF(Random Forest) ML(machine learning) 알고리즘과 CNN(Convolutional Neural Network) DL(deep learning) 신경망의 하이브리드화를 포함하여 여러 가지 장점을 보여주며 개별 CNN 및 RF 모델에 비해 정확도가 향상될 수 있다. 또한, RS 데이터를 사용함으로써 토양 질감에 영향을 미치는 다양한 변수를 적절한 규모로 비용을 절감하면서 추출할 수 있다.

도면의 간단한 설명

[0027] 도 1은 본 발명의 일실시예에 있어서 컴퓨터 장치의 내부 구성의 일례를 설명하기 위한 블록도이다.

도 2는 토양의 물리적 특성 예측을 위한 대상 지역 예시를 도시한 것이다.

도 3은 대상 지역 토지피복(Landcover) 분류 예시를 도시한 것이다.

도 4는 토양 질감 통계적 요약을 나타낸 것이다.

도 5는 이상값이 제거된 토양 질감 통계적 요약을 나타낸 것이다.

도 6은 토양 질감의 특성에 영향을 미치는 매개변수를 나타낸 것이다.

도 7은 RS(Remote sensing) 매개변수를 나타낸 것이다.

도 8은 RF(Random Forest) 알고리즘의 아키텍처를 도시한 것이다.

도 9는 CNN(Convolutional Neural Network) 모델의 아키텍처를 도시한 것이다.

도 10은 하이브리드 CNN-RF 네트워크 아키텍처를 도시한 것이다.

도 11은 토양 질감 특성의 공간 예측을 위한 워크플로우를 도시한 것이다.

발명을 실시하기 위한 구체적인 내용

[0028] 이하, 본 발명의 실시예를 첨부된 도면을 참조하여 상세하게 설명한다.

[0030] 본 발명의 실시예들은 토양의 물리적 특성을 예측하는 기술에 관한 것이다.

[0031] 본 명세서에서 구체적으로 개시되는 것들을 포함하는 실시예들은 RF와 CNN이라는 두 가지 알고리즘의 조합을 통해 토양 질감 특성의 공간 예측 정확도를 높일 수 있다.

[0032] 본 발명의 실시예들에 따른 토양 특성 예측 장치는 적어도 하나의 컴퓨터 장치에 의해 구현될 수 있으며, 본 발명의 실시예들에 따른 토양 특성 예측 방법은 토양 특성 예측 장치에 포함되는 적어도 하나의 컴퓨터 장치를 통해 수행될 수 있다. 이때, 컴퓨터 장치에는 본 발명의 일실시예에 따른 컴퓨터 프로그램이 설치 및 구동될 수 있고, 컴퓨터 장치는 구동된 컴퓨터 프로그램의 제어에 따라 본 발명의 실시예들에 따른 토양 특성 예측 방법을 수행할 수 있다. 상술한 컴퓨터 프로그램은 컴퓨터 장치와 결합되어 토양 특성 예측 방법을 컴퓨터에 실행시키기 위해 컴퓨터 판독 가능한 기록매체에 저장될 수 있다.

[0033] 도 1은 본 발명의 일실시예에 따른 컴퓨터 장치의 예를 도시한 블록도이다. 예를 들어, 본 발명의 실시예들에 따른 토양 특성 예측 장치는 도 1을 통해 도시된 컴퓨터 장치(100)에 의해 구현될 수 있다.

[0034] 도 1에 도시된 바와 같이 컴퓨터 장치(100)는 본 발명의 실시예들에 따른 토양 특성 예측 방법을 실행하기 위한 구성요소로서, 메모리(110), 프로세서(120), 통신 인터페이스(130) 그리고 입출력 인터페이스(140)를 포함할 수 있다.

[0035] 메모리(110)는 컴퓨터에서 판독 가능한 기록매체로서, RAM(random access memory), ROM(read only memory) 및 디스크 드라이브와 같은 비소멸성 대용량 기록장치(permanent mass storage device)를 포함할 수 있다. 여기서 ROM과 디스크 드라이브와 같은 비소멸성 대용량 기록장치는 메모리(110)와는 구분되는 별도의 영구 저장 장치로서 컴퓨터 장치(100)에 포함될 수도 있다. 또한, 메모리(110)에는 운영체제와 적어도 하나의 프로그램 코드가 저장될 수 있다. 이러한 소프트웨어 구성요소들은 메모리(110)와는 별도의 컴퓨터에서 판독 가능한 기록매체로부터 메모리(110)로 로딩될 수 있다. 이러한 별도의 컴퓨터에서 판독 가능한 기록매체는 플로피 드라이브, 디스크, 테이프, DVD/CD-ROM 드라이브, 메모리 카드 등의 컴퓨터에서 판독 가능한 기록매체를 포함할 수 있다. 다른 실시예에서 소프트웨어 구성요소들은 컴퓨터에서 판독 가능한 기록매체가 아닌 통신 인터페이스(130)를 통해 메모리(110)에 로딩될 수도 있다. 예를 들어, 소프트웨어 구성요소들은 네트워크(160)를 통해 수신되는 파일들에 의해 설치되는 컴퓨터 프로그램에 기반하여 컴퓨터 장치(100)의 메모리(110)에 로딩될 수 있다.

[0036] 프로세서(120)는 기본적인 산술, 로직 및 입출력 연산을 수행함으로써, 컴퓨터 프로그램의 명령을 처리하도록 구성될 수 있다. 명령은 메모리(110) 또는 통신 인터페이스(130)에 의해 프로세서(120)로 제공될 수 있다. 예를 들어, 프로세서(120)는 메모리(110)와 같은 기록 장치에 저장된 프로그램 코드에 따라 수신되는 명령을 실행하도록 구성될 수 있다.

[0037] 통신 인터페이스(130)는 네트워크(160)를 통해 컴퓨터 장치(100)가 다른 장치와 서로 통신하기 위한 기능을 제

공할 수 있다. 일례로, 컴퓨터 장치(100)의 프로세서(120)가 메모리(110)와 같은 기록 장치에 저장된 프로그램 코드에 따라 생성한 요청이나 명령, 데이터, 파일 등이 통신 인터페이스(130)의 제어에 따라 네트워크(160)를 통해 다른 장치들로 전달될 수 있다. 역으로, 다른 장치로부터의 신호나 명령, 데이터, 파일 등이 네트워크(160)를 거쳐 컴퓨터 장치(100)의 통신 인터페이스(130)를 통해 컴퓨터 장치(100)로 수신될 수 있다. 통신 인터페이스(130)를 통해 수신된 신호나 명령, 데이터 등은 프로세서(120)나 메모리(110)로 전달될 수 있고, 파일 등은 컴퓨터 장치(100)가 더 포함할 수 있는 저장 매체(상술한 영구 저장 장치)로 저장될 수 있다.

[0038] 통신 방식은 제한되지 않으며, 네트워크(160)가 포함할 수 있는 통신망(일례로, 이동통신망, 유선 인터넷, 무선 인터넷, 방송망 등)을 활용하는 통신 방식뿐만 아니라 기기들 간의 근거리 유선/무선 통신 역시 포함될 수 있다. 예를 들어, 네트워크(160)는, PAN(personal area network), LAN(local area network), CAN(campus area network), MAN(metropolitan area network), WAN(wide area network), BBN(broadband network), 인터넷 등의 네트워크 중 하나 이상의 임의의 네트워크를 포함할 수 있다. 또한, 네트워크(160)는 버스 네트워크, 스타 네트워크, 링 네트워크, 메쉬 네트워크, 스타-버스 네트워크, 트리 또는 계층적(hierarchical) 네트워크 등을 포함하는 네트워크 토폴로지 중 임의의 하나 이상을 포함할 수 있으나, 이에 제한되지 않는다.

[0039] 입출력 인터페이스(140)는 입출력 장치(150)와의 인터페이스를 위한 수단일 수 있다. 예를 들어, 입력 장치는 마이크, 키보드, 카메라 또는 마우스 등의 장치를, 그리고 출력 장치는 디스플레이, 스피커와 같은 장치를 포함할 수 있다. 다른 예로 입출력 인터페이스(140)는 터치스크린과 같이 입력과 출력을 위한 기능이 하나로 통합된 장치와의 인터페이스를 위한 수단일 수도 있다. 입출력 장치(150)는 컴퓨터 장치(100)와 하나의 장치로 구성될 수도 있다.

[0040] 또한, 다른 실시예들에서 컴퓨터 장치(100)는 도 1의 구성요소들보다 더 적은 혹은 더 많은 구성요소들을 포함할 수도 있다. 그러나, 대부분의 종래기술적 구성요소들을 명확하게 도시할 필요성은 없다. 예를 들어, 컴퓨터 장치(100)는 상술한 입출력 장치(150) 중 적어도 일부를 포함하도록 구현되거나 또는 트랜시버(transceiver), 카메라, 각종 센서, 데이터베이스 등과 같은 다른 구성요소들을 더 포함할 수도 있다.

[0041] 이하에서는 토양의 물리적 특성 예측을 위한 지리공간 인공지능과 위성영상의 융합 기술의 구체적인 실시예를 설명하기로 한다.

[0042] 본 발명은 RF ML 알고리즘과 CNN DL 신경망의 하이브리드화를 통해 높은 정확도로 토양의 물리적 특성을 예측할 수 있다.

[0043] 예측 대상 영역(Study Area)

[0044] 도 2는 토양의 물리적 특성 예측을 위한 대상 지역 위치, 토양 시료 분포, 기상 관측소를 나타내고 있다.

[0045] 대상 지역의 일례로, 이란 Golestan 지역은 위도 36° 56'에서 37° 35'까지, 경도 54° 58'에서 55° 42'까지이다. 중부와 북서부를 포함한 대부분의 지역은 밀 재배에 전념하고 있으며 남부와 북동부는 주로 방목에 사용된다. 이 지역의 가장 높은 고도와 가장 낮은 고도는 각각 해발 1722m와 0m이다. 대상 지역의 연평균 강우량과 대기 온도는 각각 456mm와 21° C이다.

[0046] 본 발명은 대상 지역이 제기하는 독특한 과제에 초점을 맞추어 방법과 데이터 소스의 융합을 사용하여 토양 물리적 특성을 예측하는 것을 목표로 한다. 예측의 목표 공간 해상도는 30m이며, 이는 예측 지도(predictive map)를 생성하는 것을 목표로 하는 규모를 반영한다.

[0047] 토양 샘플(Soil Samples)

[0048] 토양 샘플 데이터는 점토(clay), 모래(sand), 실트(silt)의 세 가지 속성을 포함하여 이란 수자원 및 토양 연구소에서 수집한 317개의 샘플(0-30cm)로 구성된다. 샘플링은 그리드 샘플링 방법을 사용하여 수행되며, 각 그리드는 $1km^2$ 면적을 커버하고 GPS(위성 항법 시스템)를 사용하여 토양 샘플의 정확한 좌표를 결정한다.

[0049] 도 3은 대상 지역의 토지피복(Landcover) 분류를 나타내고 있다. 총 317개의 토양 샘플이 다양한 토지 피복 계층에 분포되어 있다. 구체적으로 샘플의 73%는 농경지에, 13%는 범위 토지에, 9%는 덮개가 없는 평야에, 3%는 주거 지역에, 1%는 산림에, 1%는 수역에 속한다. 모든 토양 샘플 중 약 75%는 200m 이하의 고도에 위치하며 나머지 25%는 200m 이상의 고도에 위치한다.

[0050] 비중계법(hydrometer method)을 이용하여 모래, 실트, 점토를 포함한 토양 질감 특성을 분석하며, 도 4의 테이블은 토양 질감 특성의 최소값, 최대값, 평균값, 표준편차값을 나타내고 있다.

[0051] 도 5의 테이블은 Theil-Sen 회귀분석을 이용하여 자료집에서 이상치를 제거한 후 토양 질감의 유형별 통계자료를 요약한 것이다. 317개의 토양 샘플로 구성된 원래 자료집은 모래, 실트, 점토에 대해 각각 179개, 144개, 155개의 샘플로 축소된다.

[0052] 환경변수(Environmental Parameters)

[0053] 이전 연구, 전문가 의견 및 대상 지역의 특정 조건을 기반으로 세 그룹의 환경 매개변수가 사용된다. 여기에는 지구 관측 위성 중 하나인 Landsat-8의 밴드 1(B1)~밴드 5(B5) 및 밴드 7(B7), 밝기 지수(BI: Brightness Index), 착색 지수(CI: Coloration Index), 점토 지수(CLI: Clay Index), 식생 강화 지수(EVI: Enhanced Vegetation Index), 지표면 온도(LST: Land Surface Temperature), 색상 지수(HI: Hue Index), 정규 식생 지수(NDVI: Normalized Difference Vegetation Index), 적색 지수(RI: Redness Index) 및 포화 지수(SI: Saturation Index)와 같은 RS 변수가 포함될 수 있다.

[0054] 기온, 강우량과 같은 기후 변수도 포함될 수 있으며, 지형 변수에는 측면(aspect), 고도(elevation), 경사(slope), 지속 방사선(DR: duration radiation), 다중 해상도 바닥 평탄도(MRVBF: multi-resolution index of valley bottom flatness), 다중 해상도 리지탑 평탄도(MRRTF: multi-resolution ridgetop flatness index), 지형 습윤도(TWI: topographic wetness index)가 포함될 수 있다.

[0055] 본 발명에서는 종속 변수가 대상 변수로 취급되는 토양 질감 특성을 나타내며, 독립 변수는 도 6의 테이블에 도시한 바와 같이 다양한 환경 변수를 포함할 수 있다.

[0056] (1) RS 매개변수

[0057] 도 7의 테이블에 나열된 바와 같이, Landsat 8 위성 이미지에서 14개의 RS 매개변수를 추출할 수 있다. 사용된 RS 이미지는 2020년 1월 1일부터 12월 30일 사이에 수집된 것이다. 이미지 위치는 Landsat 글로벌 참조 시스템의 경로 162, 34행, 경로 162, 35행, 경로 163, 34행에 해당한다. Landsat 8 OLI 센서 이미지는 Google 어스에서 방사선 및 기하학적으로 보정되어 WGS84-Zone 40 N에 투영될 수 있다.

[0058] (2) 지형 매개변수

[0059] 본 발명에서 사용된 지형 매개변수는 Google 어스 엔진 시스템과 ArcGIS 10.8 및 SAGA 8.2.1 소프트웨어를 사용하여 공간 해상도 30X30m의 SRTM(Shuttle Radar Topography Mission) 디지털 지형 모델에서 추출된다. 지형 매개변수에는 측면(aspect), 표고(elevation), 경사(slope), DR(Duration Radiation), MRVBF(Multi-Resolution index of Valley Bottom Flatness), MRRTF(Multi-Resolution Ridgetop Flatness index) 및 TWI(Topographic Wetness Index)가 포함될 수 있다.

[0060] TWI는 수학식 1을 통해 계산될 수 있다.

[0061] [수학식 1]

$$\mathrm{TWI} = \ln \frac{A_s}{\tan \beta}$$

[0062]

[0063] 여기서, A_s는 집수구역 지수(catchment area index)이고, β는 경사각(slope angle)을 의미한다.

[0064] (3) 기후 매개변수

[0065] 본 발명에서 사용한 기후 매개변수는 도 2에 도시한 바와 같이 Golestan 지역 10개 기상관측소의 연평균(2014~2020) 자료에서 획득할 수 있다. ArcGIS 10.8 소프트웨어를 사용하여 자료에 다양한 보간법을 적용할 수 있다. RMSE 지수를 기반으로 대기 온도 및 강우량 지도를 생성하는 데 가장 정확한 방법은 로컬 다항식 방법이다.

[0066] 토양의 질감에 영향을 미치는 매개변수 분석

[0067] RF 알고리즘에서 토양 점토에 가장 영향력 있는 매개변수는 MRVBF로, 이는 다양한 크기와 경사의 계곡 바닥을 식별하여 해당 지역에 대한 더 나은 설명을 제공한다. MRVBF에는 MRVBF와 유사하게 점토 함량이 고지대에서 평지로 갈수록 점토 함량이 증가하는 점토와 직접적으로 관련된 지역의 위치에 대한 정보가 포함되어 있다. 토양 모래의 경우 가장 효과적인 매개변수는 토지 피복 유형과 지역 환경 조건에 의해 흡수되는 태양 에너지의 양에

따라 달라지는 LST이다. 모래와 농업 지역은 토지 피복 구조로 인해 가장 높은 온도를 흡수한다. 모래에 영향을 미치는 두 번째 매개변수는 B5로, 물 댐이 있는 지역에서 모래의 양이 가장 많고, B5 반사의 양은 가장 낮다. 이는 물이 근적외선을 가장 많이 흡수하기 때문에, B5 매개변수와 대상 지역의 모래 양 사이에 반대 관계가 발생하기 때문이다. 연구 결과는 B7이 토양 미사를 예측하는 데 가장 중요한 환경 변수임을 나타낸다. SWIR(단파장 적외선) 밴드는 토양 질감 특성, 특히 미사를 예측하고 추정하는 데 중요한 역할을 한다. 더욱이, 미사는 SWIR 밴드의 반사 및 흡수 강도에 영향을 미치는 요인 중 하나이다. 특징 중요도를 결정한 후 RS 매개변수가 각 토양 질감 특성을 예측하는 데 가장 중요한 기여를 하는 것으로 나타난다. RS는 이러한 매개변수에 적절한 공간적 및 시간적 정확성을 제공한다. 선택된 RS 매개변수에는 RI, B7, SI, CI, B5, NDVI, CLI 등 7개의 매개변수가 포함되며, B7과 NDVI가 가장 큰 영향을 미친다. NDVI는 스펙트럼 측정에서 대기와 토양 배경의 영향을 줄이는 가장 널리 사용되는 식생 지수 중 하나이다. SI, CI, RI는 Landsat 8의 가시광선 3개 대역(B2, B3, B4)(도 7의 테이블)에서 추출된 매개변수로 토양 특성 예측에 중요한 영향을 미친다. 이러한 매개변수는 짧은 주기성, 우수한 공간 분해능 및 적용 범위, 가시광선, B5 및 SWIR을 포함한 광범위한 스펙트럼 범위와 같은 장점을 갖춘 Landsat 8 데이터에서 얻는다. RS 변수를 통합하면 예측 정확도가 향상될 수 있다.

[0068] 예측 모델(Prediction Models)

[0069] (1) RF 알고리즘

[0070] RF 알고리즘은 더 높은 정확도를 달성하기 위해 여러 의사 결정 트리의 예측 결과를 결합하는 앙상블 학습 기법이다. RF 알고리즘은 다양한 분야에서 널리 사용되고 있으며, 분류, 회귀 및 비지도 학습 문제를 해결하는 데 탁월한 성능을 보이고 있다.

[0071] RF에서 트리 예측기 집합(set of tree predictors) $h(x;\theta_k),\ k=1,\ldots,K$ 가 사용되며, 여기서 x는 관측치의 입력 벡터(변수)를 나타내고 Θ_k는 독립적이고 동일하게 분포된 무작위 벡터(random vector)이다. 원본 데이터 세트를 대체하는 각 Θ_k는 회귀 트리에 맞춰진다. 각 트리의 각 노드에 대해 작은 입력 변수 집합이 무작위로 고려된다. 트리 분할 기준은 가장 낮은 지니(Gini) 지수를 가진 입력 변수를 선택하는 것을 기반으로 한다. 마지막으로, 회귀 문제에서 RF 예측의 출력은 전체 의사 결정 트리 집합의 가중치가 없는 평균이다.

[0072] [수학식 2]

$$h(x)=\left(\frac{1}{k}\right)\sum_{k=1}^{K}h(x;\theta_k)$$

[0073]

[0074] RF 알고리즘의 전체 흐름도는 도 8과 같다.

[0075] (2) CNN

[0076] CNN은 살아있는 유기체의 시각 인식 메커니즘에서 영감을 받은 DL에 대한 아키텍처이다. CNN은 컨볼루션, 최대 풀링, 드롭아웃, 연결, 완전 연결 등 여러 계층으로 구성된다. 컨볼루션 계층에는 입력 데이터로부터 서로 다른 특징을 계산하는 여러 커널이 포함된다. 최상위 풀링 계층은 각 영역의 최대 특징 수를 다음 계층에 입력으로 보내어 행렬의 차원을 줄이고 과적합을 피할 수 있다. 드롭아웃은 과적합을 방지하는 또 다른 방법이다.

[0077] 컨볼루션 계층의 출력 C_j는 수학식 3을 통해 계산할 수 있다.

[0078] [수학식 3]

$$C_j=f\left(b+\sum_{i}^{n}\mathrm{conv1D}(W_{ij},\ x_i)\right),\ j=1,2,\ldots,k$$

[0079]

[0080] 여기서, x_i는 CNN 네트워크의 입력 벡터의 i번째 특징이고, W_{ij}는 x_i와 바이패스 b를 가진 컨볼루션 계층의 j번째 커널 사이의 가중치이며, k와 n은 각각 컨볼루션 계층에 대한 입력 벡터의 커널 수와 특징 수이다. 활성화 함수 f는 그 중에서도 시그모이드, tanh 또는 ReLU일 수 있다.

[0081] 이러한 CNN 아키텍처는 도 9와 같다.

[0082] (3) CNN-RF

[0083] 본 발명에서는 DL과 ML의 하이브리드 네트워크를 사용함으로써 CNN 네트워크와 RF 알고리즘의 기능을 모두 활용하여 토양 질감의 공간 예측에서 더 높은 성능을 달성하고 이러한 독립형 모델의 한계를 극복할 수 있다.

[0084] 하이브리드 CNN-RF 네트워크 아키텍처는 도 10에 도시한 바와 같다. 입력 행렬은 mXn 구조를 가정하며, 여기서 m은 토양 샘플의 양을 나타내고 n은 각 토양 질감 속성에 영향을 미치는 매개변수의 수를 나타낸다. 입력 정보는 먼저 CNN 모델의 히든 레이어를 통해 처리되며, 입력 데이터 세트에서 공간 패턴과 컨텍스트 정보를 포함한 관련 특징을 추출한다. 그런 다음, 이러한 특징은 회귀 분석을 위해 RF 알고리즘에 입력된다. 마지막으로 출력 계층은 예측 값을 반환한다.

[0085] 모델 평가(Models Evaluation)

[0086] 모델의 효율성은 RMSE(Root Mean Squared Error), MSE(Mean Squared Error), 결정 계수(coefficient of determination)(R^2)의 세 가지 지표를 사용하여 평가된다. MSE 및 RMSE 값이 낮을수록 모델 정확도가 높음을 나타낸다. R^2는 데이터와 회귀 모델 간의 적합도를 나타낸다. R^2 값의 범위는 0에서 1이며, 1에 가까운 값은 더 나은 모델 성능을 나타낸다.

[0087] [수학식 4]

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

[0088]

[0089] [수학식 5]

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

[0090]

[0091] [수학식 6]

$$R^2 = 1 - \frac{MSE}{\frac{1}{n}\sum_{i=1}^{n}(y_i - \overline{y})^2}$$

[0092]

[0093] 여기서, y_i는 측정값, $\hat{y}_i$는 예측값, $\overline{y}$는 실제 값의 평균, n은 관측치의 수를 나타낸다.

[0094] 통계 지표 간의 관계를 표시하고 토양 특성을 예측하는 데 있어 모델 성능의 차이를 시각화하는 또 다른 효과적인 방법은 테일러 다이어그램을 사용하는 것이다. 테일러 다이어그램은 상관 관계에 대해 예측값과 관측값 사이의 일치 정도와 표준 편차 오차를 보여준다. 또한, 박스 플롯(box plot)을 사용하여 범위의 최소값과 최대값, 상위 사분위수와 하위 사분위수, 예측값과 실제 데이터의 중간값을 비교한다. 이 값 집합은 데이터 집합의 분포에 대한 간략한 요약을 제공한다.

[0095] K-폴드 교차 검증(K-Fold Cross-Validation)

[0096] 교차 검증은 ML 및 DL 모델의 성능을 강력하고 편향되지 않은 방식으로 평가하는 데 사용되는 기술이다. 10배 교차 검증에서는 데이터 세트가 거의 동일한 크기의 10배로 분할된다. 데이터 세트는 무작위로 10개의 부분 집합으로 분할되며, 각 부분 집합에는 동일한 수의 샘플이 포함된다. 이는 폴드에 걸쳐 데이터 분포가 전체 데이터 세트를 대표하도록 보장한다. 그런 다음 교차 검증 프로세스는 반복적으로 수행되며, 각 폴드는 테스트 세트로 사용되고 나머지 9배는 모델 학습에 사용된다. 성능 지표는 모델의 예측을 기반으로 각 반복에 대해 계산된다.

[0097] 토양의 질감 예측을 위한 워크플로우(Workflow for Soil Texture Prediction)

[0098] 도 11은 본 발명에 따른 토양 질감 특성의 공간 예측을 위한 워크플로우를 도시한 것이다.

[0099] 도 11을 참조하면, 첫 번째 단계는 위성 이미지에서 추출한 매개변수와 대상 지역에서 수집한 데이터를 사용하여 공간 데이터베이스를 만드는 것이다. 두 번째 단계에서는 추출된 매개변수를 RF 알고리즘을 통해 특징 중요도를 결정하기 위해 독립적인 데이터로 사용한다. 세 번째 단계에서는 RF, CNN, CNN-RF 알고리즘을 사용하여 토양 질감 특성을 모델링한다. 네 번째 단계에서는 모델을 사용하여 토양 질감 특성에 대한 예측 지도를 생성한다. 마지막으로, MSE, RMSE, R^2, 박스 플롯, 테일러 다이어그램의 5가지 지표를 사용하여 모델 평가를 수행한다.

[0100] 하이브리드 CNN-RF 모델은 개별 CNN 및 RF 모델에 비해 더 높은 정밀도를 나타낸다. 모델링 프로세스 초기에 사용된 컨볼루션 레이어를 통해 입력 데이터 특징을 추출하고 구성할 수 있다. 반대로, CNN이 최종 회귀 결정을 위해 완전 연결 계층을 사용하면 종종 과적합이 발생한다. 결과적으로, RF 알고리즘을 통합하면 결과의 정확도가 향상될 수 있다.

[0101] 따라서, 본 실시예들은 토양 질감 특성의 공간적 예측을 위한 하이브리드 CNN-RF 모델을 제공할 수 있다. 토양 질감에 영향을 미치는 지표를 준비하는 데 적절한 공간적, 시간적 정확성으로 인해 위성 이미지가 사용된다. (1) RF 알고리즘은 MRVBF, LST 및 B7을 각각 점토, 모래 및 미사에 대한 가장 효과적인 매개변수로 식별한다. (2) 유효 매개변수 중 RS 변수가 모델링 입력에 가장 큰 기여를 한다. 구체적으로, NDVI, B7, SI, B5, CI, RI 및 CLI는 토양 질감에 영향을 미치는 중요한 RS 매개변수라 할 수 있다. (3) 하이브리드 CNN-RF 모델은 토양의 질감 특성 예측에서 가장 높은 정확도를 나타낸다. (4) 모래, 미사 및 점토는 MSE 평가 지표를 기반으로 더 높은 정확도를 나타낸다.

[0102] 하이브리드 CNN-RF 모델을 통해 생성된 예측 지도는 농업 관리, 토양 침식 모니터링 및 관개에 도움이 될 수 있다. 실시예에 따라서는 (1) 모델링 정확도를 향상시키기 위해 RF 알고리즘 대신 메타 휴리스틱 알고리즘(meta-heuristic algorithm)을 활용할 수 있다. (2) 회색조 동시 발생 행렬(gray-level cooccurrence matrix)을 이용하여 조사 대상 지역의 균질성, 대비, 상이성, 엔트로피 등의 변수를 추출하여 토양 성상 예측 정확도를 높일 수 있다.

[0103] 이처럼 본 발명의 실시예들에 따르면, 토양 질감의 공간 예측을 위한 RF ML 알고리즘과 CNN DL 신경망의 하이브리드화를 포함하여 여러 가지 장점을 보여주며 개별 CNN 및 RF 모델에 비해 정확도가 향상될 수 있다. 또한, RS 데이터를 사용함으로써 토양 질감에 영향을 미치는 다양한 변수를 적절한 규모로 비용을 절감하면서 추출할 수 있다.

[0104] 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, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPGA(field programmable gate 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 어플리케이션을 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.

[0105] 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, 컴퓨터 저장 매체 또는 장치에 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.

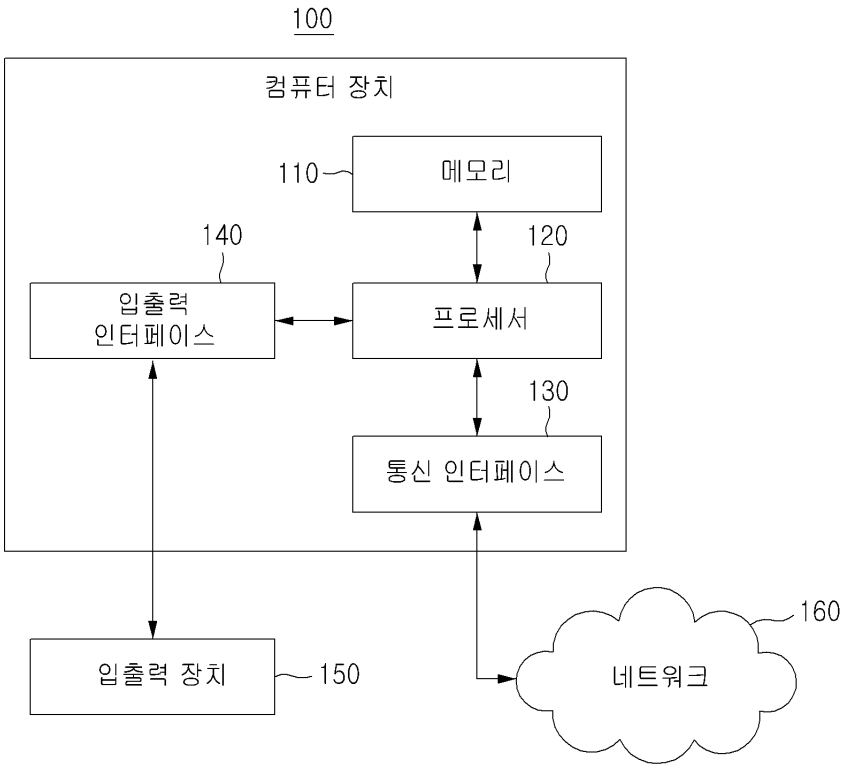
[0106] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 이때, 매체는 컴퓨터로 실행 가능한 프로그램을 계속 저장하거나, 실행 또는 다운로드를 위해 임시 저장하는 것일 수도 있다. 또한, 매체는 단일 또는 수 개의 하드웨어가 결합된 형태의 다양한 기록수단 또는 저장수단일 수 있는데, 어떤 컴퓨터 시스템에 직접 접속되는 매체에 한정되지 않고, 네트워크 상에 분산 존재하는 것일 수도 있다. 매체의 예시로는, 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체, CD-ROM 및 DVD와 같은 광기록 매체, 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical medium), 및 ROM, RAM, 플래시 메모리 등을 포함하여 프로그램 명령어가 저장되도록 구성된 것이 있을 수 있다. 또한, 다른 매체의 예시로, 어플리케이션을 유통하는 앱 스토어나 기타 다양한 소프트웨어를 공급 내지 유통하는 사이트, 서버 등에서 관리하는 기록매체 내지 저장매체도 들 수 있다.

[0107] 이상과 같이 실시예들이 비록 한정된 실시예와 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.

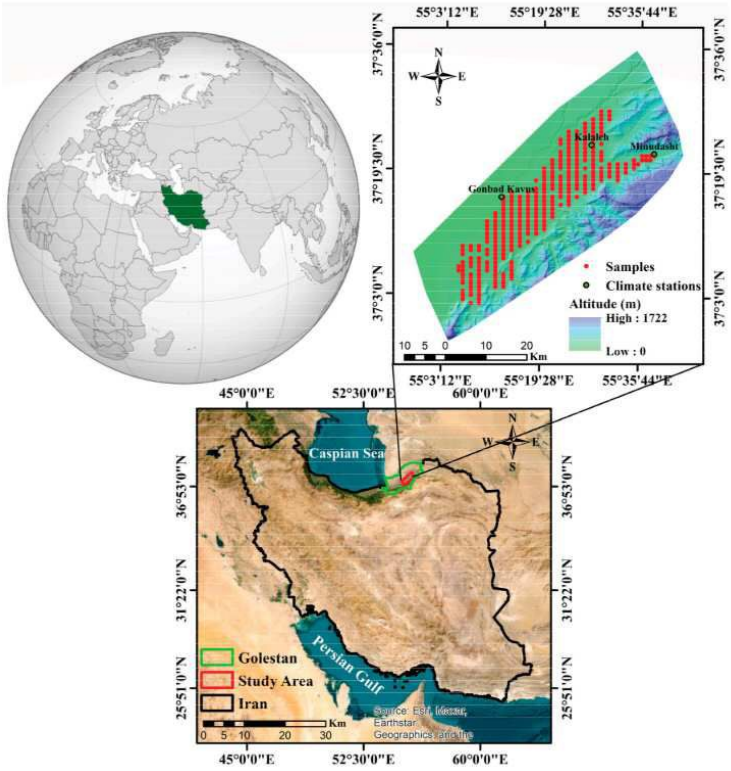
[0108] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

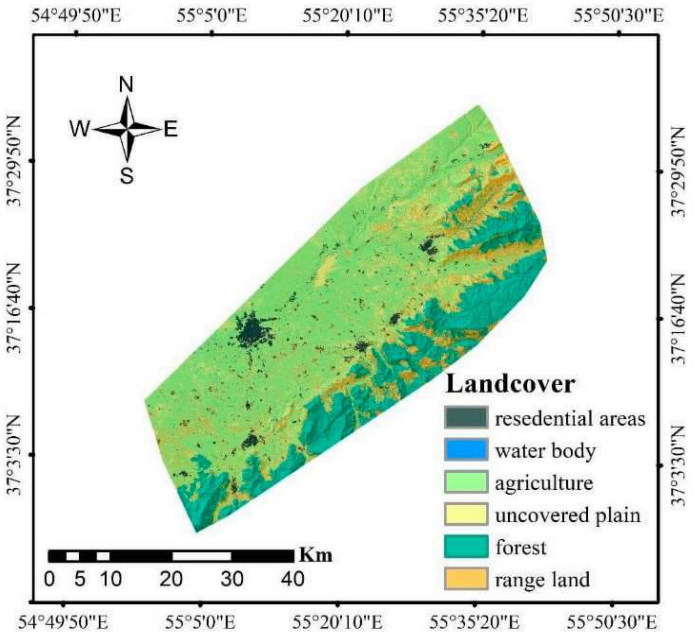
도면1



도면2


55°3'12"E
55°19'28"E
55°35'44"E
37°36'0"N
37°19'30"N
37°3'0"N
Kalaleh
Minudasht
Gonbad Kavus
Samples
Climate stations
Altitude (m)
High : 1722
Low : 0
10 5 0 10 20 Km
45°0'0"E
52°30'0"E
60°0'0"E
36°53'0"N
31°22'0"N
25°51'0"N
Caspian Sea
Persian Gulf
Golestan
Study Area
Iran
0 5 10 20 30 Km

도면3


54°49'50"E
55°5'0"E
55°20'10"E
55°35'20"E
55°50'30"E
37°29'50"N
37°16'40"N
37°3'30"N
N
W
E
S
Landcover
resedential areas
water body
agriculture
uncovered plain
forest
range land
0 5 10 20 30 40 Km

도면4

Soil Texture	Clay (%)	Silt (%)	Sand (%)
Minimum	0	0	0
Maximum	44	80	58
Mean	22.322	64.457	12.867
Standard deviation	6.920	9.249	9.115

도면5

Soil Texture	Clay (%)	Silt (%)	Sand (%)
Minimum	12	50	4
Maximum	36	76	26
Mean	22.182	5.74	11.056
Standard deviation	4.199	4.567	3.705

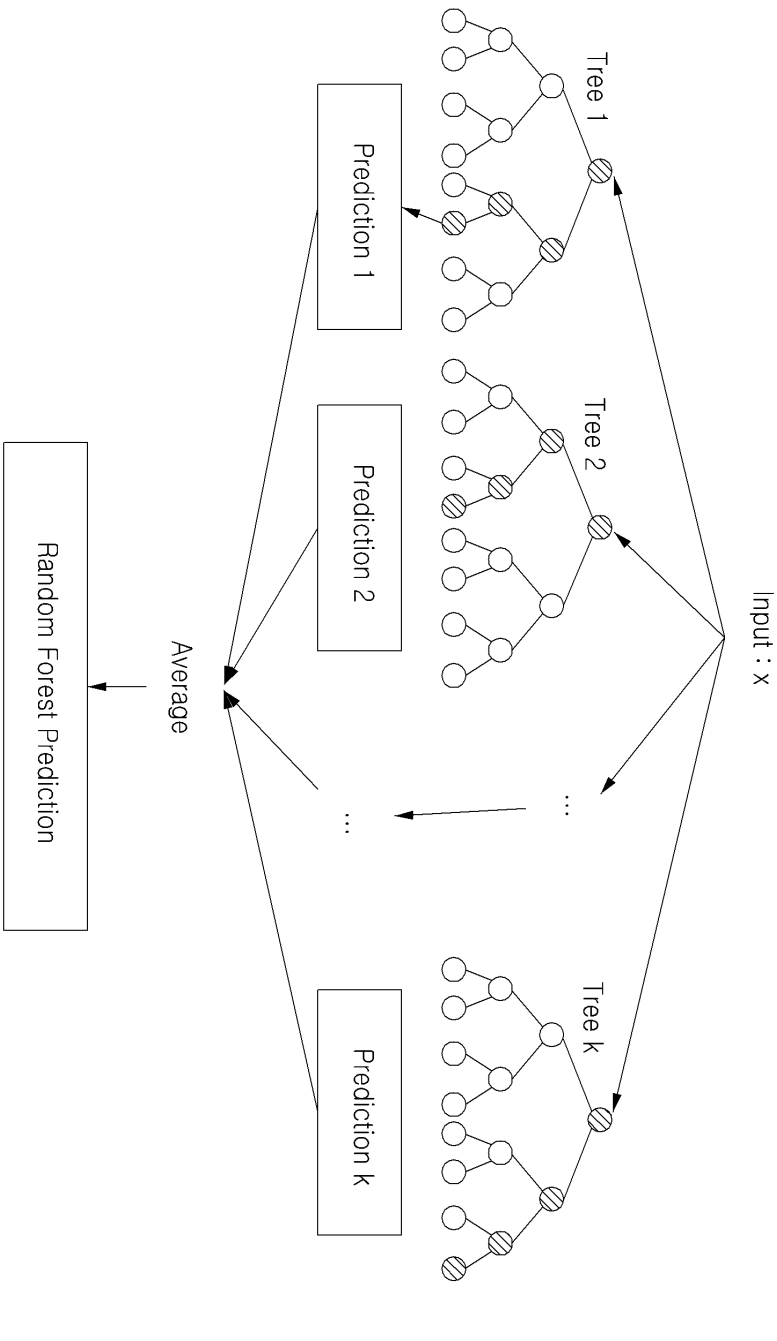
도면6

Soil Texture	Effective Parameters	Number of Parameters
Clay	NDVI, Elevation, B7, B5, B1, B2, B3, B4, MRRTF, MRVBF, Rainfall, SI, CI, LST, Temp, Aspect, RI, TWI	18
Silt	NDVI, Elevation, B7, B5, B3, B4, MRRTF, MRVBF, SI, BI, CLI, CI, Slope, EVI, DR, Aspect, RI, TWI	18
Sand	NDVI, Elevation, B7, B5, B1, B2, B3, B4, Rainfall, SI, BI, CLI, MRRTF, MRVBF, CI, Slope, LST, DR	18

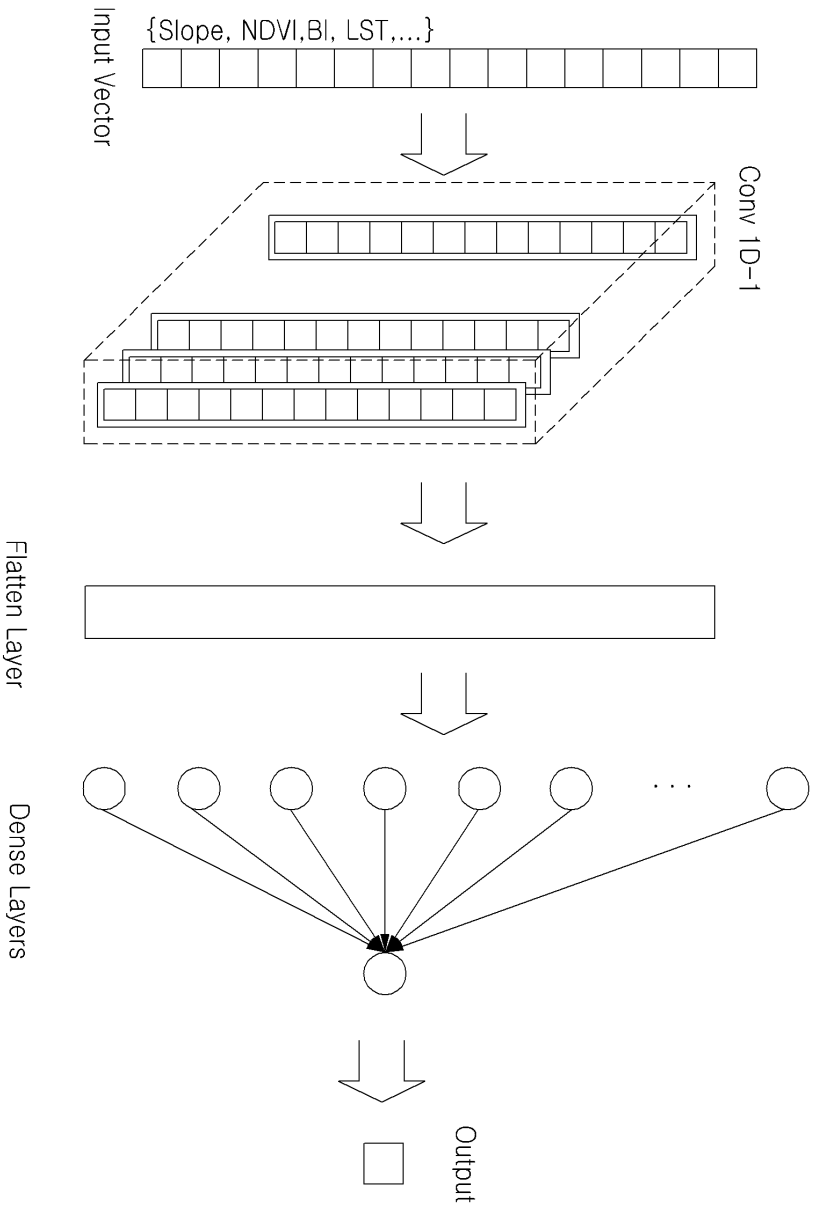
도면7

Covariate Name	Definition
Coastal aerosol (B1)	0.43–0.45 μm
Blue (B2)	0.45–0.51 μm
Green (B3)	0.53–0.59 μm
Red (B4)	0.64–0.67 μm
Near-infrared (B5)	0.85–0.88 μm
Short-wave infrared-2 (B7)	2.11–2.29 μm
Brightness Index (BI)	$(B3^2 + B4^2)^{0.5}$
Clay Index (CLI)	B6/B7
Coloration Index (CI)	(B4 − B3)/(B4 + B3)
Enhanced Vegetation Index (EVI)	$2.5 \times \left(\frac{B5-B4}{B5+(6\times B4)-(7.5\times B2)+1}\right)$
Land Surface Temperature (LST)	
Normalized Difference Vegetation Index (NDVI)	(B5 − B4)/(B5 + B4)
Redness Index (RI)	$(B4^2)/\ (B2 \times (B3^3))$
Saturation Index (SI)	(B4 − B2)/(B4 + B2)

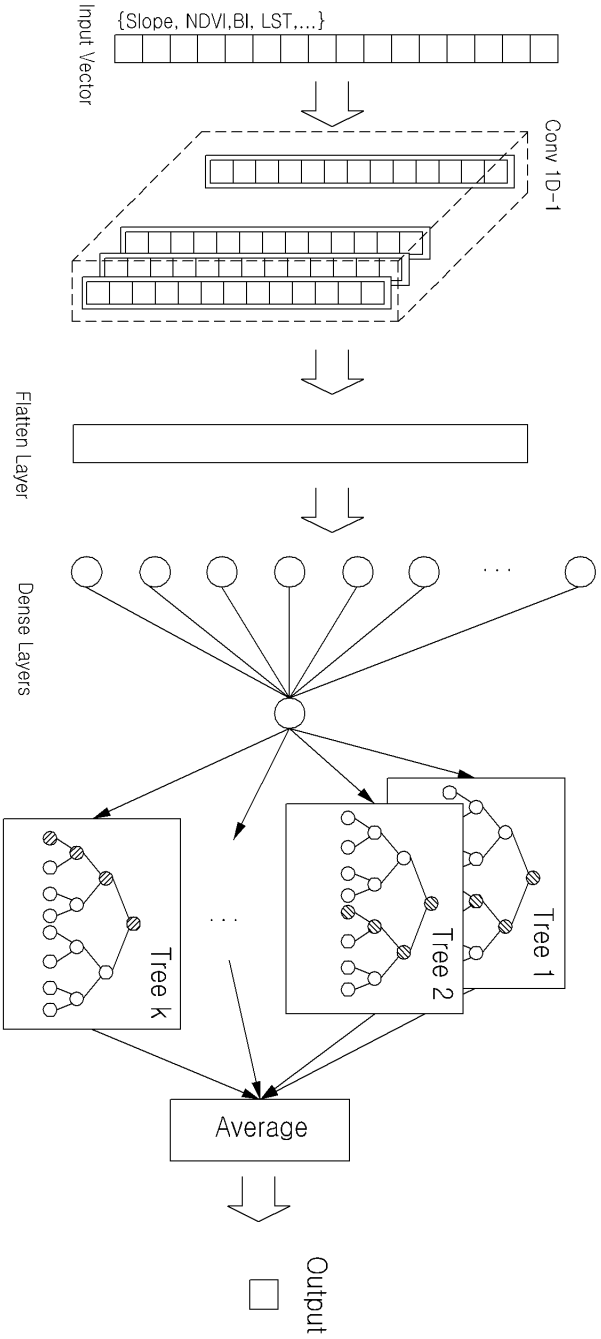
도면8


Input : x
Tree 1
Tree 2
...
Tree k
Prediction 1
Prediction 2
...
Prediction k
Average
Random Forest Prediction

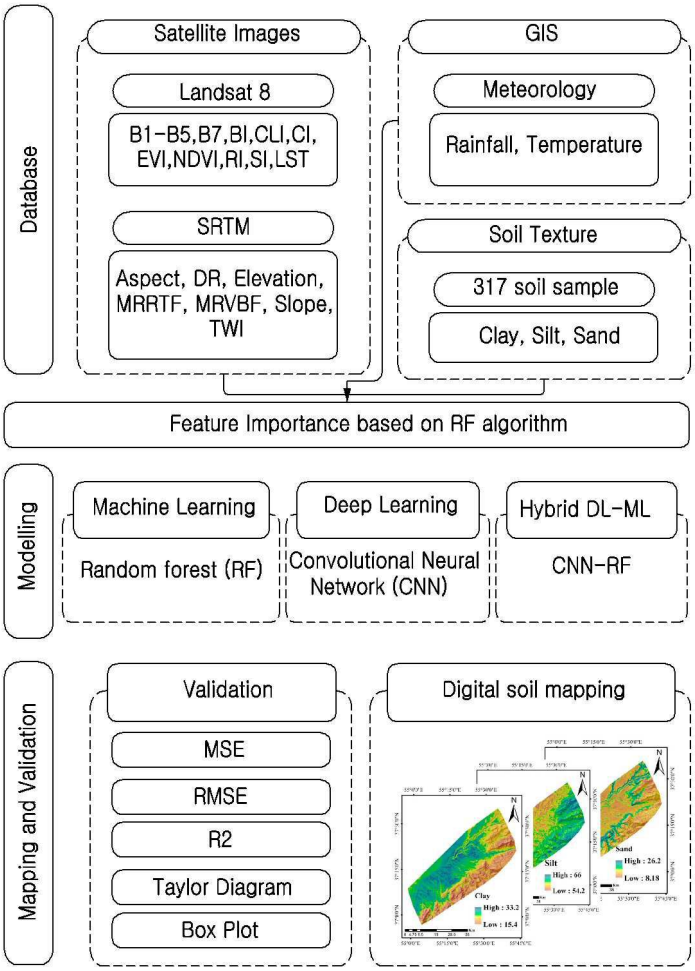
도면9


Input Vector
{Slope, NDVI,BI, LST,...}
Conv 1D-1
Flatten Layer
Dense Layers
Output

도면10


Input Vector
{Slope, NDVI,BI, LST,...}
Conv 1D-1
Flatten Layer
Dense Layers
Tree k
Tree 2
Tree 1
Average
Output

도면11


Database
Satellite Images
Landsat 8
B1-B5,B7,BI,CLI,CI,
EVI,NDVI,RI,SI,LST
SRTM
Aspect, DR, Elevation,
MRRTF, MRVBF, Slope,
TWI
GIS
Meteorology
Rainfall, Temperature
Soil Texture
317 soil sample
Clay, Silt, Sand
Feature Importance based on RF algorithm
Modelling
Machine Learning
Random forest (RF)
Deep Learning
Convolutional Neural
Network (CNN)
Hybrid DL-ML
CNN-RF
Mapping and Validation
Validation
MSE
RMSE
R2
Taylor Diagram
Box Plot
Digital soil mapping
Clay
High : 33.2
Low : 15.4
Silt
High : 66
Low : 54.2
Sand
High : 26.2
Low : 8.18