(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2023년04월13일
(11) 등록번호 10-2520763
(24) 등록일자 2023년04월07일

(51) 국제특허분류(Int. Cl.)
G01N 30/86 (2006.01) *G01N 30/72* (2006.01)
G06N 3/08 (2023.01) *G16C 20/70* (2019.01)
(52) CPC특허분류
G01N 30/8693 (2013.01)
G01N 30/7233 (2013.01)
(21) 출원번호 10-2022-0053215
(22) 출원일자 2022년04월29일
심사청구일자 2022년04월29일

(73) 특허권자
울산과학기술원
울산광역시 울주군 언양읍 유니스트길 50
(72) 발명자
조경화
울산광역시 울주군 언양읍 유니스트길 50, 113동 406호 (울산과학기술원)
윤다은
울산광역시 울주군 언양읍 유니스트길 50, 113동 405호 (울산과학기술원)
(뒷면에 계속)
(74) 대리인
김종선, 이형석

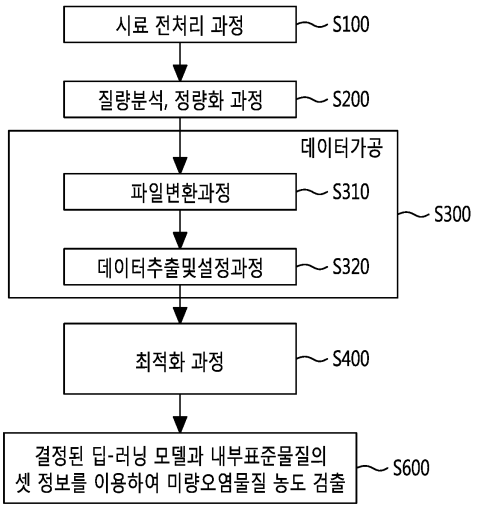
전체 청구항 수 : 총 8 항 심사관 : 김기환

(54) 발명의 명칭 **딥-러닝 모델을 이용한 미량오염물질 정량화 장치 및 방법**

(57) 요 약

본 발명은 미량오염물질(MP)의 분석 및 정량화 시 투입되는 내부표준물질(SIL-IS)의 개수를 최적화하고 분석 데이터를 기반으로 하여 더 정확하게 미량오염물질(MP)을 검출할 수 있는 정량화 장치 및 그 방법을 제안한다. 본 발명의 정량화 장치는, 정량화된 미량오염물질(MP)의 변환 전 데이터를 활용 가능한 데이터로 변환하는 데이터 변환부와, 상기 변환된 데이터로부터 미량오염물질(MP), 내부표준물질(SIL-IS), 체크 스탠다드(Check standard, CS) 정보, 자연유기물질(NOM) 정보를 추출하여 입력 서브셋(subset) 데이터로 가공하는 입력 데이터 가공부, 상기 가공된 입력 서브셋 데이터를 입력받고, 내부표준물질(SIL-IS)의 개수 및 종류, 딥-러닝 모형의 종류를 최적화하는 최적화부, 그리고 최적화된 내부표준물질(SIL-IS)의 개수 및 종류와, 딥-러닝 모형을 조합하여 미량오염물질(MP)의 농도를 정량화하는 모델 학습부를 포함하여 구성된다.

대 표 도 - 도3



(52) CPC특허분류
G06N 3/08 (2023.01)
G16C 20/70 (2019.02)

(72) 발명자
백상수
경상북도 경산시 대학로 280, 소재관 508호 (대동, 영남대학교)
전준호
경상남도 양산시 물금읍 신주로 64, 102동 2502호 (양산신도시현진에버빌)
강대호
경상남도 창원시 의창구 팔용로407번길 8, 1층 (팔용동)

이 발명을 지원한 국가연구개발사업

과제고유번호	1485018538
과제번호	ARQ202001348003
부처명	환경부
과제관리(전문)기관명	한국환경산업기술원
연구사업명	수생태계 건강성 확보 기술개발사업
연구과제명	신종 미량 오염물질 수생태계 유입 부하량 예측 기술 개발
기 여 율	1/1
과제수행기관명	울산과학기술원
연구기간	2020.04.22 ~ 2023.12.31

공지예외적용 : 있음

명 세 서

청구범위

청구항 1

채취된 환경 샘플에서 획득한 오비트랩 매스 데이터(Orbitrap Mass data)을 mzXML 파일로 변환하는 데이터 변환부;

상기 mzXML 파일에서 확인 가능한 미량오염물질(MP), 내부표준물질(SIL-IS), 체크 스탠다드(Check standard, CS) 정보, 자연유기물질(NOM) 정보를 가공하는 입력 데이터 가공부;

상기 가공된 입력 데이터를 다목적 최적화 프로세스를 이용하여 상기 내부표준물질(SIL-IS)의 종류와 개수를 최적화하고, 최적화된 내부표준물질(SIL-IS)의 정보에 따라 딥-러닝 모델을 최적화하는 최적화부;

상기 최적화된 내부표준물질(SIL-IS)의 개수 및 종류와, 딥-러닝 모델을 조합하여 미량오염물질(MP)의 농도를 정량화하는 모델 학습부를 포함하여 구성되는 것을 특징으로 하는, 미량오염물질 정량화 장치.

청구항 2

제 1 항에 있어서,

상기 체크 스탠다드(CS)는 측정 기기에 대해 일간, 일중의 정밀도(intra and inter-day precision)에 대한 정보를 포함하고,

상기 자연유기물질(NOM)은 수소 대 탄소(H/C) 및 산소 대 탄소(O/C)의 2가지 몰(Molar) 비율 정보를 포함하는, 미량오염물질 정량화 장치.

청구항 3

제 1 항에 있어서,

상기 다목적 최적화 프로세스,

NSGA-II(Non-Dominated Sorting Genetic Algorithm)을 이용하는, 미량오염물질 정량화 장치.

청구항 4

제 1 항에 있어서,

상기 모델 학습부는, CNN 학습모델인, 미량오염물질 정량화 장치.

청구항 5

미량오염물질(MP)의 농도를 추정하는 장치가,

채취한 환경샘플로부터 수집한 오비트랩 매스 데이터(Orbitrap Mass data)을 mzXML 파일로 변환하는 단계;

상기 변환된 mzXML 파일에서 확인 가능한 미량오염물질(MP), 내부표준물질(SIL-IS), 체크 스탠다드(Check standard, CS) 정보, 자연유기물질(NOM) 정보를 가공하는 단계;

상기 가공된 정보들을 기초로 다목적 최적화 프로세스를 이용하여 상기 내부표준물질(SIL-IS)의 종류와 개수를 최적화하고, 최적화된 내부표준물질(SIL-IS)의 정보에 따라 딥-러닝 모델을 최적화하는 단계;

상기 최적화된 내부표준물질(SIL-IS)의 개수 및 종류와, 딥-러닝 모형을 조합하여 미량오염물질(MP)의 농도를 정량화하는 단계를 포함하여 수행되는, 미량오염물질 정량화 방법.

청구항 6

제 5 항에 있어서,

상기 'mzXML' 데이터는 Raw Converter에 의해 변환하는, 미량오염물질 정량화 방법.

청구항 7

제 5 항에 있어서,

상기 다목적 최적화 프로세스는 NSGA-II(Non-Dominated Sorting Genetic Algorithm)을 이용하여 수행하는, 미량오염물질 정량화 방법.

청구항 8

제 5 항에 있어서,

상기 정량화된 미량오염물질(MP)의 농도와 상기 미량오염물질의 농도와 대응하는 내부표준물질(SIL-IS) 정보를 환경 데이터로 저장하는 단계를 더 포함하는, 미량오염물질 정량화 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 미량오염물질(Micro-Pollutants, MP)의 분석 및 정량화 시 투입되는 내부표준물질(Stable Isotope-Labeled Internal standard: SIL-IS)의 종류 및 개수를 최적화하고 분석 데이터를 기반으로 하여 더 정확하게 미량오염물질(MP)을 검출하며, 또 정량화 시 사용되는 딥-러닝(Deep- learning) 모델의 종류를 최적화하여 분석 데이터의 형태에 적합한 모델이 적용될 수 있도록 한 미량오염물질 정량화 장치 및 그 방법에 관한 것이다.

배 경 기 술

[0002] 미량오염물질(MP)이란 ng/L에서 ㎍/L 또는 이보다 더 낮은 농도로 수중에 존재하는 오염물질들을 말하며, 수중 생태계로 방출되는 미량오염물질(MP)은 공중 보건에 부정적인 영향을 미친다. 최근에는 기존 미량오염물질(MP) 외에 새로운 분석기술의 등장으로 인해 새롭게 발견되거나 기존 법령에 규제되지 않고 있는 신종 미량오염물질이 새로운 오염원으로서 생태계를 위협하고 있다.

[0003] 이와 같은 미량오염물질(MP)의 정밀 분석을 위해서는 측정된 농도가 정확한가에 대한 검증이 반드시 필요하다. 현재까지 분석 기술에서 미량오염물질(MP)을 정밀 분석할 수 있는 가장 신뢰성 있는 방법은 측정되는 물질에 대한 기준이 될 수 있는 안정적 동위원소 표지 내부표준물질(Stable Isotope-Labeled Internal standard; SIL-IS, 이하 '내부표준물질'이라 약칭하여 설명함)를 넣어서 확인하는 방법이다.

[0004] 그런데 미량오염물질(MP) 분석을 위해 사용되는 상기 내부표준물질(SIL-IS)은 수십~ 수백만원 정도의 고가 소모품이다. 따라서 분석할 시료가 많은 경우 시료마다 미량오염물질(MP)을 분석하기 위해서는 비용이 그만큼 증가하게 된다.

[0005] 종래에 이러한 내부표준물질(SIL-IS)의 사용을 배제하여 미량오염물질(MP)을 검출하는 방안이 제안된 바 있다. 즉 본 출원인이 출원하고 2022. 01. 06. 자로 공개된 10-2022-0001972호(딥러닝 기반 모델링을 통한 미량오염물질 정밀검출장치 및 방법)가 있다. 상기 기술은 합성곱 신경망(CNN)을 이용하여 대량의 자연유기물질(NOM) 데이터로부터 결과를 도출할 수 있어, 환경시료의 미량오염물질의 농도를 정확하게 측정할 수 있다.

[0006] 하지만 종래기술에서 자연유기물질(NOM)은 지역별로 특성이 상이하게 제공되기 때문에 다른 지역 등에 적용하는데 문제가 있다. 또 상기 종래기술은 측정 데이터(Retention time, mass-to-charge ratio)를 기반으로 내부표준물질(SIL-IS)과 유사한 특징을 보이는 부분을 추출하여 미량오염물질(MP)을 정량화하였는데, 이때 측정 기계의 오차(intra-/inter-day precision)를 전혀 반영하기 않았다. 또한 CNN 모델 선정 시 사용자가 직접 모델을 선정하여 학습을 진행하였기 때문에 그 선정된 학습모델이 적합한 모델인지 판정할 수 있는 기준이 마련되어 있지 않았다. 따라서 종래기술에서는 그만큼 의미 있거나 신뢰성 있는 정보를 확보할 수 없는 문제점이 있었다.

선행기술문헌

특허문헌

[0007] (특허문헌 0001) 한국공개특허 10-2022-0001972(2022. 01. 06)

발명의 내용

해결하려는 과제

[0008] 본 발명의 목적은 상기와 같은 문제점을 해결하기 위하여 안출된 것으로서, 미량오염물질(MP)의 검출에 사용되는 내부표준물질(SIL-IS)의 개수를 최적화하는 미량오염물질 정량화 장치 및 방법을 제공하는 것이다.

[0009] 본 발명의 다른 목적은 미량오염물질(MP)을 분석 데이터를 기반으로 검출하는 미량오염물질 정량화 장치 및 방법을 제공하는 것이다.

[0010] 본 발명의 또 다른 목적은 미량오염물질(MP)의 분석 데이터에 적합한 학습 모델의 종류를 최적화하여 미량오염물질 정량화 장치에 적합한 최적의 학습모델을 제공하는 것이다.

[0011] 본 발명의 기술적 과제들은 이상에서 언급한 기술적 과제로 제한되지 않으며, 언급되지 않은 또 다른 기술적 과제들은 아래의 기재로부터 당업자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

[0012] 이와 같은 목적을 달성하기 위한 본 발명의 일 실시 예에 따른 미량오염물질 정량화 장치는, 채취된 환경 샘플에서 획득한 오비트랩 매스 데이터(Orbitrap Mass data)을 mzXML 파일로 변환하는 데이터 변환부; 상기 mzXML 파일에서 확인 가능한 미량오염물질(MP), 내부표준물질(SIL-IS), 체크 스탠다드(Check standard, CS) 정보, 자연유기물질(NOM) 정보를 가공하는 입력 데이터 가공부; 상기 가공된 입력 데이터를 다목적 최적화 프로세스를 이용하여 상기 내부표준물질(SIL-IS)의 종류와 개수를 최적화하고, 최적화된 내부표준물질(SIL-IS)의 정보에 따라 딥-러닝 모델을 최적화하는 최적화부; 상기 최적화된 내부표준물질(SIL-IS)의 개수 및 종류와, 딥-러닝 모델을 조합하여 미량오염물질(MP)의 농도를 정량화하는 모델 학습부를 포함하여 구성되는 것을 특징으로 한다.

상기 체크 스탠다드(CS)는 측정 기기에 대해 일간, 일중의 정밀도(intra and inter-day precision)에 대한 정보를 포함하고, 상기 자연유기물질(NOM)은 수소 대 탄소(H/C) 및 산소 대 탄소(O/C)의 2가지 몰(Molar) 비율 정보를 포함한다.

상기 다목적 최적화 프로세스, NSGA-II(Non-Dominated Sorting Genetic Algorithm)을 이용한다.

상기 모델 학습부는, CNN 학습모델이다.

본 발명의 다른 실시 예에 따른 미량오염물질 정량화 방법은, 미량오염물질(MP)의 농도를 추정하는 장치가, 채취한 환경샘플로부터 수집한 오비트랩 매스 데이터(Orbitrap Mass data)을 mzXML 파일로 변환하는 단계; 상기 변환된 mzXML 파일에서 확인 가능한 미량오염물질(MP), 내부표준물질(SIL-IS), 체크 스탠다드(Check standard, CS) 정보, 자연유기물질(NOM) 정보를 가공하는 단계; 상기 가공된 정보들을 기초로 다목적 최적화 프로세스를 이용하여 상기 내부표준물질(SIL-IS)의 종류와 개수를 최적화하고, 최적화된 내부표준물질(SIL-IS)의 정보에 따라 딥-러닝 모델을 최적화하는 단계; 및 상기 최적화된 내부표준물질(SIL-IS)의 개수 및 종류와, 딥-러닝 모형을 조합하여 미량오염물질(MP)의 농도를 정량화하는 단계를 포함하여 수행되는 것을 특징으로 한다.

상기 'mzXML' 데이터는 Raw Converter에 의해 변환한다.

상기 다목적 최적화 프로세스는 NSGA-II(Non-Dominated Sorting Genetic Algorithm)을 이용하여 수행한다.

상기 정량화된 미량오염물질(MP)의 농도와 상기 미량오염물질의 농도와 대응하는 내부표준물질(SIL-IS) 정보를 환경 데이터로 저장하는 단계를 더 포함한다.

[0013] 삭제

[0014] 삭제

[0015] 삭제

[0016] 삭제

[0017] 삭제

[0018] 삭제

[0019] 삭제

발명의 효과

[0020] 이와 같은 본 발명에 따르면, 다목적 최적화 프로세스를 통해 미량오염물질(MP)의 검출 시 필요한 내부표준물질(SIL-IS)의 종류와 개수, 그리고 딥-러닝 모델의 모형을 최적화할 수 있다. 따라서 내부표준물질(SIL-IS)의 개수를 줄일 수 있어, 표적분석법의 정확도를 높이기 위해 고가의 내부표준물질(SIL-IS)을 사용했던 종래 기술보다 비용을 절감할 수 있는 효과가 있다.

[0021] 본 발명에 따르면, 최적화된 딥-러닝 모형과 내부표준물질(SIL-IS)을 조합하여 딥-러닝 모델을 훈련하고 검증하는 과정을 수행하기 때문에, 검출 결과의 정확도가 향상되는 효과가 있다.

[0022] 본 발명에 따르면, 자연유기물질(NOM) 정보를 활용하여 지역별로 다양하게 나타나는 자연유기물질(NOM)을 반영할 수 있어 전 지역적으로 적용하여 미량오염물질(MP)의 농도를 산출할 수 있는 효과가 있다.

[0023] 본 발명에 따르면, 최적화된 딥-러닝 모델과 내부표준물질 셋 정보를 이용함으로써 미량오염물질(MP)의 검출시간은 물론 작업자의 노동력이 단축되는 효과도 기대할 수 있다.

도면의 간단한 설명

[0024] 도 1은 본 발명의 바람직한 실시 예에 따른 미량오염물질 정량화 장치의 구성도이다.

도 2는 도 1의 입력 데이터 가공부가 처리하는 데이터의 예를 보인 도면이다.

도 3은 본 발명의 바람직한 실시 예에 따른 미량오염물질 정량화 장치를 이용하여 미량오염물질의 농도를 검출하는 과정을 설명하는 흐름도이다.

발명을 실시하기 위한 구체적인 내용

[0025] 본 발명은 다양한 변환을 가할 수 있고 여러 가지 실시 예를 가질 수 있는 바, 특정 실시 예들을 도면에 예시하고 상세하게 설명하고자 한다. 그러나, 이는 본 발명의 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변환, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다. 본 발명을 설명함에 있어서 관련된 공지 기술에 대한 구체적인 설명이 본 발명의 요지를 흐릴 수 있다고 판단되는 경우 그 상세한 설명을 생략한다.

[0026] 제1, 제2 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 상기 구성요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다.

[0027] 본 발명에서 사용한 용어는 단지 특정한 실시 예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 출원에서, "포함하다" 또는 "가지다" 등의 용어는 명세서상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.

[0028] 공간적으로 상대적인 용어인 아래(below, beneath, lower), 위(above, upper) 등은 도면에 도시되어 있는 바와 같이 하나의 소자 또는 구성 요소들과 다른 소자 또는 구성 요소들과의 상관 관계를 용이하게 기술하기 위해 사용될 수 있다. 공간적으로 상대적인 용어는 도면에 도시되어 있는 방향에 더하여 사용시 또는 동작시 소자의 서

로 다른 방향을 포함하는 용어로 이해되어야 한다. 예를 들면, 도면에 도시되어 있는 소자를 뒤집을 경우, 다른 소자의 아래(below, beneath)로 기술된 소자는 다른 소자의 위(above, upper)에 놓여질 수 있다. 따라서, 예시적인 용어인 아래는 아래와 위의 방향을 모두 포함할 수 있다. 소자는 다른 방향으로도 배향될 수 있고, 이에 따라 공간적으로 상대적인 용어들은 배향에 따라 해석될 수 있다.

[0029] 본 발명에서 사용되는 “부” 또는 “부분” 등의 일부분을 나타내는 표현은 해당 구성요소가 특정 기능을 포함할 수 있는 장치, 특정 기능을 포함할 수 있는 소프트웨어, 또는 특정 기능을 포함할 수 있는 장치 및 소프트웨어의 결합을 나타낼 수 있음을 의미하나, 꼭 표현된 기능에 한정된다고 할 수는 없으며, 이는 본 발명의 보다 전반적인 이해를 돕기 위해서 제공된 것일 뿐, 본 발명이 속하는 분야에서 통상적인 지식을 가진 자라면 이러한 기재로부터 다양한 수정 및 변형이 가능하다.

[0030] 또한, 본 발명에서 사용되는 모든 전기 신호들은 일 예시로서, 본 발명의 회로에 반전기 등을 추가적으로 구비하는 경우 이하 설명될 모든 전기 신호들의 부호가 반대로 바뀔 수 있음을 유의해야 한다. 따라서, 본 발명의 권리범위는 신호의 방향에 한정되지 않는다.

[0031] 따라서, 본 발명의 사상은 설명된 실시 예에 국한되어 정해져서는 아니되며, 후술하는 특허청구범위뿐 아니라 이 특허청구범위와 균등하거나 등가적 변형이 있는 모든 것들은 본 발명 사상의 범주에 속한다고 할 것이다.

[0032] 이하에서는 도면에 도시한 실시 예에 기초하면서 본 발명에 대하여 더욱 상세하게 설명하기로 한다.

[0033] 도 1은 본 발명의 바람직한 실시 예에 따른 미량오염물질 정량화 장치의 구성도이다.

[0034] 도 1에 도시한 바와 같이 미량오염물질 정량화 장치(10)는 분석/정량화부(100), 데이터 변환부(200), 입력데이터 가공부(300), 최적화부(400) 및 모델학습부(500)를 포함하여 구성된다.

[0035] 실시 예에서 분석/정량화부(100)는 오비트랩(Orbitrap) 기반의 분석 기기일 수 있고, 오비트랩 매스 데이터(Orbitrap Mass data)의 수집을 위하여 고해상도 질량분석(LC-HRMS) 및 미량오염물질을 정량화하는 유닛이다. 알려진 바와 같이 질량 분석은, 이온 질량 및 전하에 따라 물질의 기체 이온을 분리함으로써 물질의 화학 성분을 식별할 수 있게 한다.

[0036] 실시 예에서 데이터 변환부(200)는 상기 분석/정량화부(100)에 의해 생성된 LC-HRMS(액체 크로마토그래피-고해상도 질량분석) 데이터를 활용 가능한 mzXML 데이터로 변환하는 유닛이다. 즉 질량분석시험에서 생성되는 대용량 데이터는 작업자가 직접 조작하기 어렵기 때문에, 개방형 XML 기반의 데이터 포맷으로 변환하는 것이다.

[0037] 실시 예에서 입력 데이터 가공부(300)는 내부표준물질(SIL-IS)의 사용 개수를 최적화하기 위하여 데이터를 가공하는 유닛이다. 구체적으로 상기 데이터 변환부(200)에 의해 변환된 변환 후 데이터, 미량오염물질(MP), 내부표준 물질(SIL-IS), 측정기기의 오차를 처리하는 Check standard(CS) 정보 및 자연유기물질(NOM)에 대한 H/C, O/C 비율 정보들을 이용하여 입력 서브셋(subset)으로 가공한다. 여기서 미량오염물질(MP), 내부표준물질(SIL-IS), 그리고 Check standard는, Retention time과 mass-to-charge ratio 정보가 이용된다. 본 실시 예는 이처럼 미량오염물질(MP)의 분석 시 사용되는 측정 기기의 일간/일내 오차(Check standard)를 반영하여 측정 기기의 분석 조건과 유사하게 입력 값을 구성하였다.

[0038] 또한 질량 분석 데이터에서 확인 가능한 자연유기물질(NOM)의 정보를 활용할 수 있어, 지역별로 다양하게 나타나는 자연유기물질(NOM)을 반영할 수 있다. 따라서 지역별 특성을 반영하여 광범위한 지역에 활용할 수 있다.

[0039] 상기 입력 데이터 가공부(300)가 처리하는 데이터의 예는 도 2에 도시하였다. 도 2를 보면 질량 스펙트럼, 대상 미량오염물질(MP), 내부표준물질(SIL-IS), Check standard(CS) 및 자연유기물질(NOM) 정보들이 모두 이용되기 때문에, 미량오염물질(MP)의 측정 정확도를 향상시킬 수 있다.

[0040] 실시 예에서 최적화부(400)는 상기 입력 데이터 가공부(300)에 의해 구축된 입력 서브셋 데이터를 이용하여 내부표준물질(SIL-IS)의 종류와 개수, 딥-러닝의 모형 종류를 최적화하는 유닛이다. 딥-러닝의 모형 종류를 최적화하는 것은 다수 개의 딥-러닝 모형 중 어느 하나의 딥-러닝 모형을 선택하는 것을 의미한다. 실시 예에서 내부표준물질(SIL-IS)의 종류와 개수, 딥-러닝의 모형 종류를 최적화하는 것은 다목적 최적화(Multi-objective optimization) 프로세스를 이용하며, 특히 NSGA-II(Non-Dominated Sorting Genetic Algorithm)를 이용한다.

그리고 실시 예에서 수식 1은 딥-러닝 손실함수의 평균제곱오차(MSE, Mean Squared Error)에 따른 값을 나타낸다.

수학식 1

[0041] $$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(\hat{Y}_i - Y_i)^2$$

[0042] 그리고 딥-러닝 모델은 CNN 학습 모델일 수 있다. 따라서 CNN 학습모델에서 특징을 추출하는 계층이나 분류를 위한 계층 등에 대한 최적화도 이루어질 것이다.

[0043] 실시 예에서 모델 학습부(500)는 상기 최적화부(400)에 의해 최적화된 내부표준물질(SIL-IS)의 종류와 개수, 딥-러닝 모형을 조합하여 훈련하고, 훈련 결과로서 미량오염물질(MP) 검출과 정량화를 자동화하고 미량오염물질(MP)의 농도를 정량화하여 제공하는 유닛이다.

[0044] 이러한 모델 학습부(500)는 최적화된 내부표준물질(SIL-IS)을 입력으로 받아 미량오염물질(MP)의 농도를 검출하기 때문에, 종래보다 내부표준물질(SIL-IS)의 개수를 줄일 수 있다.

[0045] 도 3은 본 발명의 바람직한 실시 예에 따른 미량오염물질 정량화 장치를 이용하여 미량오염물질의 농도를 검출하는 과정을 설명하는 흐름도이다.

[0046] 일련의 대상 유역에서 환경 샘플이 채취되면, 채취된 환경 샘플에 포함된 불순물을 제거하고 타켓 미량오염물질을 농축하도록 전- 처리하는 과정이 수행될 수 있다(S100).

[0047] 이러한 전처리 과정이 완료되면 분석/정량화부(100)가 환경 샘플에 대한 질량분석과 정량화 과정을 수행한다(S200). 이때 질량분석은 오비트랩 질량분석기(Orbitrap Mass spectrometer)가 이용되며, 정량화는 내부표준방법이 사용될 수 있다. 상기 과정은 오비트랩 질량분석데이터를 수집하기 위한 과정이라 할 수 있다.

[0048] 다음에는 최적화를 위해 데이터를 가공하는 과정이 수행된다(S300). 실시 예는 입력 데이터 가공부(300)가 파일 변환과정(S310), 데이터 추출 및 설정과정(S320)을 포함하여 실시한다.

[0049] 상기 파일 변환과정(S310)은 데이터 변환부(200)가 상기 정량화 과정에 따라 생성된 LC-HRMS 데이터를 활용 가능한 'mzXML' 데이터로 변환하는 과정을 수행하는 과정이다. 상기 HRMS 데이터에는 RT(Retention time) 및 질량대 전하비(mass-to-charge ratio) 정보가 포함되어 있다. 상기 RT는 가스 크로마토그래피나 액체 크로마토그래피 분석에서 시료를 주입하고 나서 유출할 때까지 걸리는 시간, M/Z는 원자 질량 번호를 이온의 전하 수로 나눈 값이다.

[0050] 그리고 상기 파일 변환과정(S310)에서 mzXML 데이터로 변환은 'Raw Converter'를 이용할 수 있다.

[0051] 제 320 단계에서 상기 데이터 추출과정은 상기 mzXML 데이터로부터 대상 미량오염물질, SIL standards, check standards, and NOM 정보에 해당하는 정보들을 추출하고, 상기 데이터 설정과정은 추출된 데이터들을 딥-러닝 모형에 적용하도록 입력 서브셋 데이터로 생성하는 과정이다. 상기 check standards에는 일간/일내(intra and inter-day) 정보를 반영하여 측정 기기 오류에 대한 영향을 최소화할 수 있도록 하였다. 그리고 자연유기물질(NOM) 정보는 HRMS 분석에 영향을 미치는 매트릭스 효과가 있기 때문에, NOM 정보 역시 입력 서브셋에 포함된다. 또 상기 자연유기물질(NOM) 정보는 수소 대 탄소(H/C) 및 산소 대 탄소(O/C)의 2가지 몰(Molar) 비로 구성된다.

[0052] 상기한 과정에 의해 입력 서브셋 데이터가 생성되면, 최적화부(400)는 상기 입력 서브셋을 입력으로 하여 최적화 과정을 수행한다(S400).

[0053] 최적화 과정(S400)은, 다목적 최적화 프로세스에 적용해서 내부표준물질(SIL-IS)의 종류와 개수를 최적화하고, 딥-러닝 모형의 종류를 최적화하는 과정이라 할 수 있다. 즉 데이터 가공과정(S300)의 분석결과를 기반으로 내부표준물질(SIL-IS)과 딥-러닝 모형을 최적화하는 것이다. 최적화를 하는 이유는 종래기술에서 언급한 바와 같이 미량오염물질(MP)을 정량화하는 방법으로 내부표준물질(SIL-IS)의 농도를 추정하는 방법 대비 시료 수에 따라 증가하게 되는 내부표준물질(SIL-IS)의 사용 비용을 줄이기 위한 것이다. 이를 위해 본 실시 예는 다목적 최적화 프로세서를 사용하여 내부표준물질(SIL-IS)과 딥-러닝 모형의 종류를 최적화하였다.

[0054] 여기서 상기 다목적 최적화 프로세서의 수행 시 변수 선택의 등결과성을 감소하고 매뉴얼 선택의 단점을 극복하도록 파레토 최적집합을 결정하는 효율적인 진화 알고리즘의 하나인 NSGA-II(Non-Dominated Sorting Genetic

Algorithm)를 이용하였다.

[0055] 이러한 최적화 과정이 완료되면, 모델 학습부(500)는 최적화된 딥-러닝 모델과 내부표준물질 셋을 이용하여 미량오염물질 농도를 정량화한다(S600).

[0056] 실시 예에 따르면 최적의 딥-러닝 모델과 내부표준물질 셋이 결정되고, 결정된 모델과 내부표준물질 셋 정보를 이용하기 때문에, 분석 결과의 정확도가 향상되며, 작업자의 작업시간이나 노동력 등을 절감할 수 있다.

[0057] 이와 같이 본 발명은 대상 유역에서 채취한 수 환경 샘플을 분석하고, 분석 결과를 기초로 내부표준물질(SIL-IS)과 딥-러닝 모형을 최적화하고, 최적화된 딥-러닝에 학습시켜 미량오염물질(MP) 검출과 정량화를 자동화하여 농도를 산출함을 알 수 있다.

[0058] 그리고 상기 산출된 미량오염물질(MP)과 대응하는 최적의 내부표준물질(SIL-IS) 정보는 환경 데이터로 저장 관리하고, 추후 고해상도 질량분석 시에 대체 기법으로 이용할 수 있다.

[0059] 이상과 같이 본 발명의 도시된 실시 예를 참고하여 설명하고 있으나, 이는 예시적인 것들에 불과하며, 본 발명이 속하는 기술 분야의 통상의 지식을 가진 자라면 본 발명의 요지 및 범위에 벗어나지 않으면서도 다양한 변형, 변경 및 균등한 타 실시 예들이 가능하다는 것을 명백하게 알 수 있을 것이다. 따라서 본 발명의 진정한 기술적 보호 범위는 첨부된 청구범위의 기술적인 사상에 의해 정해져야 할 것이다.

부호의 설명

[0060] 10: 본 발명의 정량화 장치

100: 분석/정량화부

200: 데이터 변환부

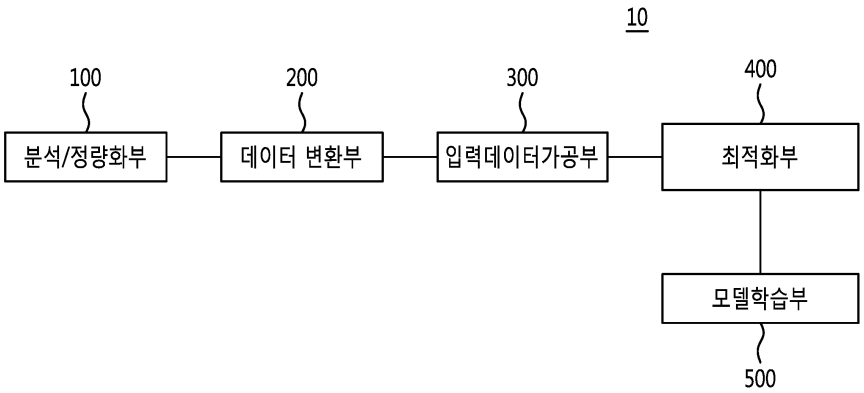
300: 입력데이터 가공부

400: 최적화부

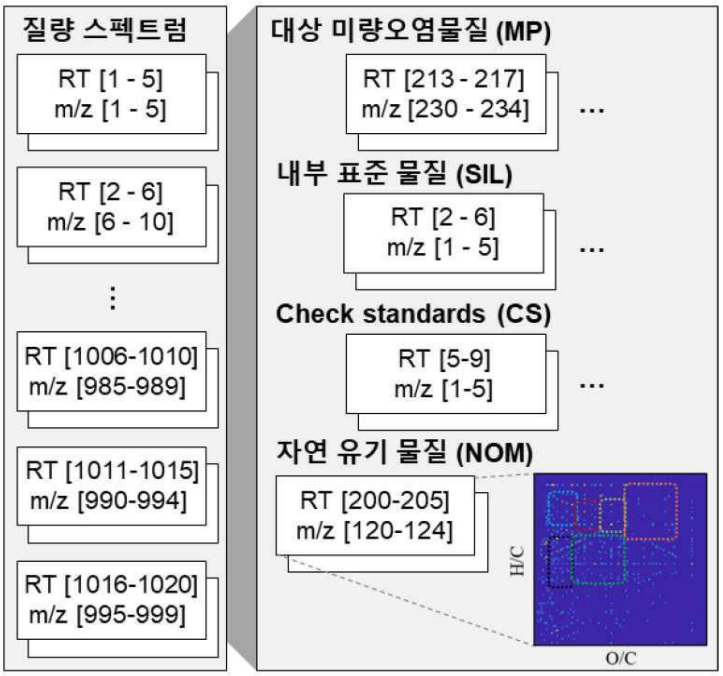
500: 모델 학습부

도면

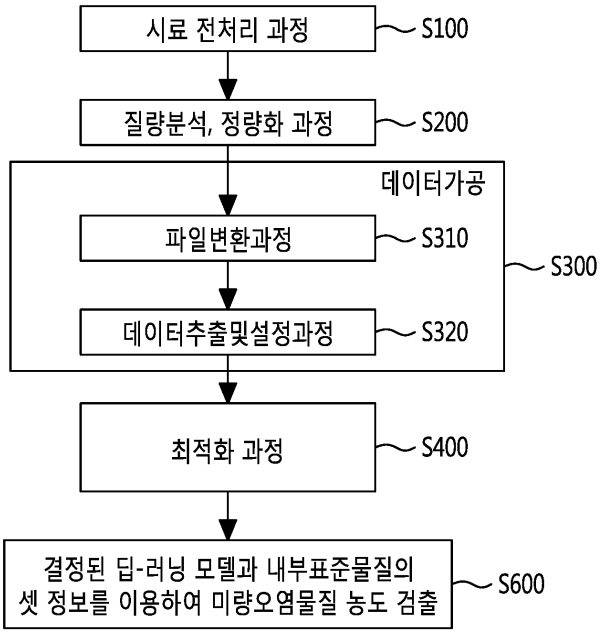
도면1



도면2



도면3



【심사관 직권보정사항】

【직권보정 1】

【보정항목】 청구범위

【보정세부항목】 청구항 1

【변경전】

채취된 환경 샘플에서 획득한 오비트랩 매스 데이터(Orbitrap Mass data)을 mzXML 파일로 변환하는 데이터 변환부;

상기 mzXML 파일에서 확인 가능한 미량오염물질(MP), 내부표준물질(SIL-IS), 체크 스탠다드(Check standard, CS) 정보, 자연유기물질(NOM) 정보를 가공하는 입력 데이터 가공부;

상기 가공된 입력 데이터를 다목적 최적화 프로세스를 이용하여 상기 내부표준물질(SIL-IS)의 종류와 개수를 최적화하고, 최적화된 내부표준물질(SIL-IS)의 정보에 따라 딥-러닝 모델을 선택하는 최적화부;

상기 최적화된 내부표준물질(SIL-IS)의 개수 및 종류와, 딥-러닝 모델을 조합하여 미량오염물질(MP)의 농도를 정량화하는 모델 학습부를 포함하여 구성되는 것을 특징으로 하는, 미량오염물질 정량화 장치.

【변경후】

채취된 환경 샘플에서 획득한 오비트랩 매스 데이터(Orbitrap Mass data)을 mzXML 파일로 변환하는 데이터 변환부;

상기 mzXML 파일에서 확인 가능한 미량오염물질(MP), 내부표준물질(SIL-IS), 체크 스탠다드(Check standard, CS) 정보, 자연유기물질(NOM) 정보를 가공하는 입력 데이터 가공부;

상기 가공된 입력 데이터를 다목적 최적화 프로세스를 이용하여 상기 내부표준물질(SIL-IS)의 종류와 개수를 최적화하고, 최적화된 내부표준물질(SIL-IS)의 정보에 따라 딥-러닝 모델을 최적화하는 최적화부;

상기 최적화된 내부표준물질(SIL-IS)의 개수 및 종류와, 딥-러닝 모델을 조합하여 미량오염물질(MP)의 농도를 정량화하는 모델 학습부를 포함하여 구성되는 것을 특징으로 하는, 미량오염물질 정량화 장치.