(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일
2026 년 7 월 16 일 (16.07.2026)

WIPO | PCT

(10) 국제공개번호
WO 2026/151326 A1

(51) 국제특허분류:
G10L 25/51 (2013.01) *G06N 3/0475* (2023.01)
G10L 25/30 (2013.01) *G06N 3/0455* (2023.01)
G10L 15/02 (2006.01) *G08B 25/00* (2006.01)
G06Q 50/26 (2012.01)

(21) 국제출원번호: PCT/KR2026/000703

(22) 국제출원일: 2026 년 1 월 13 일 (13.01.2026)

(25) 출원언어: 한국어

(26) 공개언어: 한국어

(30) 우선권정보:
10-2025-0004514 2025 년 1 월 13 일 (13.01.2025) KR

(71) 출원인: 주식회사 메타크라우드 (**METACROWD CORP.**) [KR/KR]; 03674 서울특별시 서대문구 거북골로 34, 3층 10314호 (KR).

(72) 발명자: 김형진 (**KIM, Hyungjin**); 04167 서울특별시 마포구 마포대로 33, 아-807호 (KR). 박경욱 (**PARK, Kyung-Wook**); 02522 서울특별시 동대문구 답십리로 69길 99, 1101호 (KR).

(74) 대리인: 특허법인(유한)유일하이스트 (**YUIL HIGHEST INTERNATIONAL PATENT AND LAW FIRM**); 06179 서울특별시 강남구 테헤란로 84길 15, 12층 (KR).

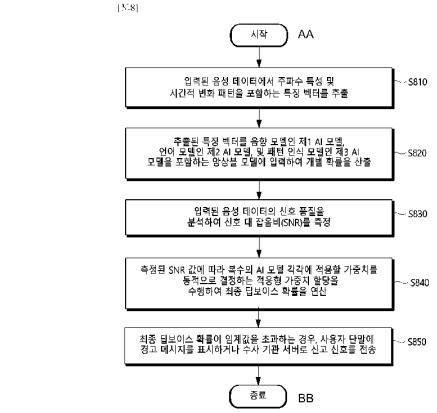
(81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG,

WO 2026/151326 A1

(54) **Title:** METHOD, DEVICE, AND SYSTEM PROVIDING ARTIFICIAL INTELLIGENCE SOLUTION FOR DETECTING DEEP VOICE GENERATED BY GENERATIVE ARTIFICIAL INTELLIGENCE AND PREVENTING RELATED INCIDENTS

(54) **발명의 명칭:** 생성형 인공지능에 의해 생성된 딥보이스를 검출하고 관련 사고를 예방하는 인공지능 솔루션 제공 방법, 장치 및 시스템

[도8]



S810 ... Extract feature vector including frequency characteristics and temporal change patterns from inputted voice data
S820 ... Input extracted feature vector into ensemble model including first AI model which is acoustic model, second AI model which is language model, and third AI model which is pattern recognition model, and calculate individual probabilities
S830 ... Analyze signal quality of inputted voice data and measure signal-to-noise ratio (SNR)
S840 ... Perform adaptive weight allocation that dynamically determines weights to be applied to each of plurality of AI models according to measured SNR value, and calculate final deep voice probability
S850 ... When final deep voice probability exceeds threshold value, display warning message on user terminal or transmit report signal to investigation agency server
AA ... Start
BB ... End

(57) **Abstract:** According to one embodiment of the present invention, provided is a system providing an artificial intelligence solution for detecting a deep voice generated by generative artificial intelligence and preventing incidents, the system comprising: a memory for storing instructions; and one or more processors for executing the instructions, wherein the processors are configured to: extract a feature vector including frequency characteristics and temporal change patterns from inputted voice data; input the extracted feature vector into an ensemble artificial intelligence model including a first AI model which is an acoustic model, a second AI model which is a language model, and a third AI model which is a pattern recognition model, so as to calculate respective individual deep voice

[다음 쪽 계속]

NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UY, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 역내 권리의 보호를 위하여): ARIPO (BW, CV, GH, GM, KE, LR, LS, MU, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 유라시아 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 유럽 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

공개:

— 국제조사보고서와 함께 (조약 제21조(3))

probabilities; measure signal quality of the inputted voice data to calculate a signal-to-noise ratio (SNR); calculate a final deep voice probability through an adaptive weight allocation algorithm that varies, in real time, weights to be applied to each AI model on the basis of the calculated SNR value; and generate a deep voice suspicion alert if the final deep voice probability exceeds a threshold value.

(57) 요약서: 본 발명의 일 형태에 따르면, 생성형 인공지능에 의해 생성된 딥보이스를 검출하는 시스템에 있어서, 인스트럭션들을 저장하는 메모리; 및 상기 인스트럭션들을 실행하는 하나 이상의 프로세서를 포함하고, 상기 프로세서는, 입력된 음성 데이터에서 주파수 특성 및 시간적 변화 패턴을 포함하는 특징 벡터를 추출하고, 상기 추출된 특징 벡터를 음향 모델인 제1 AI 모델, 언어 모델인 제2 AI 모델, 및 패턴 인식 모델인 제3 AI 모델을 포함하는 앙상블 인공지능 모델에 입력하여 각각의 개별 딥보이스 확률을 산출하며, 상기 입력된 음성 데이터의 신호 품질(Signal Quality)을 측정하여 신호 대 잡음비(SNR)를 산출하고, 상기 산출된 SNR 값에 기초하여 상기 각 AI 모델에 적용될 가중치를 실시간으로 변동시키는 적응형 가중치 할당(Adaptive Weight Allocation) 알고리즘을 통해 최종 딥보이스 확률을 연산하고, 상기 최종 딥보이스 확률이 임계값을 초과하는 경우, 딥보이스 의심 경보를 생성하도록 구성된, 딥보이스 검출 및 사고 예방 인공지능 솔루션 제공 시스템이 제공된다.