



(51) International Patent Classification:
A61B 5/00 (2006.01) G16H 40/67 (2018.01)
A61B 5/16 (2006.01)

(21) International Application Number:
PCT/IL2024/050611

(22) International Filing Date:
23 June 2024 (23.06.2024)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/523,090 25 June 2023 (25.06.2023) US

(71) Applicant: **RAMOT AT TEL-AVIV UNIVERSITY LTD.** [IL/IL]; P.O. Box 39296, 6139201 Tel Aviv (IL).

(72) Inventors: **KOBO, Oren**; c/o Ramot at Tel-Aviv University Ltd., P.O. Box 39296, 6139201 TEL AVIV (IL).
SCHONBERG, Tom; c/o Ramot at Tel-Aviv Univer-

sity Ltd., P.O. Box 39296, 6139201 TEL AVIV (IL).
MELTZER-ASSCHER, Aya Tova; c/o Ramot at Tel-Aviv University Ltd., P.O. Box 39296, 6139201 TEL AVIV (IL).
BERANT, Jonathan; c/o Ramot at Tel-Aviv University Ltd., P.O. Box 39296, 6139201 TEL AVIV (IL).

(74) Agent: **MINTOS, Moshe**; GASSNER MINTOS, P.O.B. 3393, 6103301 Tel Aviv (IL).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): 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, 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,

(54) Title: CLASSIFICATION OF DEPRESSION TENDENCY FROM GAZE PATTERNS

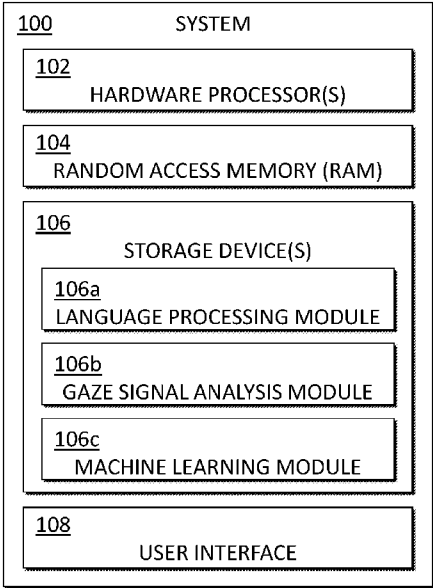


FIG. 1

(57) Abstract: A computer-implemented method comprising: receiving a plurality of gaze vectors associated with a cohort of subjects, wherein each of the gaze vectors represents tracking of a point-of-gaze of one of the subjects while performing a cognitive task; and at a training stage, training a machine learning model on a training dataset comprising: (i) all of the gaze vectors, and (ii) annotations indicating a depression score of each of the subjects, to obtain a trained machine learning model configured to issue a prediction of depression in a target subject, based on target gaze vectors obtained from gaze tracking recording of the target subject while performing the cognitive task.

TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS,
ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (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).

Published:

— *with international search report (Art. 21(3))*

CLASSIFICATION OF DEPRESSION TENDENCY FROM GAZE PATTERNS**CROSS-REFERENCE TO RELATED APPLICATIONS**

[0001] This application claims priority from U.S. Application Ser. No. 63/523,090, filed June 25, 2023, entitled “CLASSIFICATION OF DEPRESSION TENDENCY FROM GAZE PATTERNS,” the contents of which are hereby incorporated herein by reference in their entirety.

FIELD OF THE INVENTION

[0002] This invention relates to the field of machine learning.

BACKGROUND

[0003] Millions of people around the world suffer from low or negative mental states or moods, which can lead to depression. According to the World Health Organization (WHO), the proportion of the global population suffering from depression in 2015 was estimated to be 4.4%. Currently, the main diagnostic tools to assess negative mood tendencies are questionnaires, such as the PHQ-9 and Beck’s Depression Inventory (BDI); patient self-reporting; practitioner clinical observation; or any combination of the above. However, all these diagnostics are prone to variance and subjectivity. In addition, the symptoms of depression often vary between individuals, and may be with other medical conditions. Another problem with current diagnosis schemes is that none of them incorporates understanding of the underlying biological pathology, or offers an indication about the relationship between symptom clusters and biology. These problems highlight the importance—both applicative and theoretical—of developing an objective, validated tool to allow the assessment of a subject’s emotional state and tendency to depression.

[0004] In recent years, the field known as computational psychiatry has emerged, harnessing data-driven computational tools to improve psychiatric research. Computational biomarkers are specifically considered to be able to assist in the classification of mental conditions, while improving theoretical understanding of the disorders. Such biomarkers contrast with traditional mental health measurements, which usually rely on psychological questionnaires, written self-report tests, and interviews and

direct observation of behavior, all of which present a range of subjective biases. Thus, computational modeling can be used to provide classification, offer a better understanding of disorders, and help determine treatment protocols based on high-dimensional data.

[0005] Eye-tracking (ET) is a technology used to continuously measure eye movements and gaze positions, while participants are performing various tasks. This method is often used to model cognitive processing and, specifically, gauge attention. Eye-tracking often uses measurements that are calculated from the unprocessed time-series signals of gaze coordinates, such as dwell time for predefined regions of interest (ROI) on a screen. Consequently, ET offers a non-invasive technique to collect objective and precise data that can be used as a surrogate for complex cognitive processes, making it an invaluable source of data for many computational cognitive modeling challenges. ET research has shown that different psychiatric and neurological conditions manifest in abnormalities in eye movement patterns. More recently, studies have exemplified the potential of using these data for classification of mental conditions.

[0006] One potential advantage of eye-tracking over other measures is that it generates a continuous signal reflecting allocation of attention over time, while multiple stimuli compete for attention. This feature makes ET particularly useful in stimuli with spatial dependencies and continuous processing, such as reading sentences. Accordingly, in psycholinguistic research, ET has been used to investigate various aspects of sentence processing, such as corrections and reanalysis, timing of processing of different information types, and the role of anticipation in parsing. While linguistic research mostly focuses on underlying processes at the group mean level, some studies focus on individual differences and on correlating personal attributes with gaze tracking data during cognitive processing.

[0007] Currently, the standard approach to analyzing eye-tracking data during reading is to extract common aggregation features for analysis. Such analysis usually includes specific regions-of-interest (ROIs) for which these measures are computed, containing one word or phrase within a longer sentence. Some common features are the time spent upon first entering an ROI, the total time spent within an ROI, and the proportion of gaze regressions following the first time entering the ROI. Thus, while the entire gaze sequence in a trial is recorded and available for analysis, in practice, only a very small portion of it is usually utilized, and even that portion undergoes extreme dimensionality

reduction. In contrast, machine learning (ML) techniques offer the ability to analyze data with high dimensionality, and discover latent patterns, which may assist in detecting various mental health conditions. These techniques often rely on large amounts of data, and can spare or reduce the need for dimensionality reduction. Thus, ML techniques can enable us to utilize the raw gaze patterns without explicitly defining measures or ROIs, allowing the model to take into consideration underlying factors in processing that are unrepresented by common manual feature extraction methods. Moreover, viewing patterns of sentences, which are processed incrementally, should benefit from ML tools applied on the entire time-series ET data collected during the performance of the cognitive task, even more than the frequently used images viewing task.

[0008] The foregoing examples of the related art and limitations related therewith are intended to be illustrative and not exclusive. Other limitations of the related art will become apparent to those of skill in the art upon a reading of the specification and a study of the figures.

SUMMARY OF THE INVENTION

[0009] The following embodiments and aspects thereof are described and illustrated in conjunction with systems, tools and methods which are meant to be exemplary and illustrative, not limiting in scope.

[0010] There is provided, in an embodiment, a system comprising: at least one hardware processor; and a non-transitory computer-readable storage medium having stored thereon program instructions, the program instructions executable by the at least one hardware processor to: receive a plurality of gaze vectors associated with a cohort of subjects, wherein each of the gaze vectors represents tracking of a point-of-gaze of one of the subjects while performing a cognitive task, and at a training stage, train a machine learning model on a training dataset comprising: (i) all of the gaze vectors, and (ii) annotations indicating a depression score of each of the subjects, to obtain a trained machine learning model configured to issue a prediction of depression in a target subject, based on target gaze vectors obtained from gaze tracking recording of the target subject while performing the cognitive task.

[0011] There is also provided, in an embodiment, a computer-implemented method comprising: receiving a plurality of gaze vectors associated with a cohort of subjects, wherein each of the gaze vectors represents tracking of a point-of-gaze of one of the

subjects while performing a cognitive task; and at a training stage, training a machine learning model on a training dataset comprising: (i) all of the gaze vectors, and (ii) annotations indicating a depression score of each of the subjects, to obtain a trained machine learning model configured to issue a prediction of depression in a target subject, based on target gaze vectors obtained from gaze tracking recording of the target subject while performing the cognitive task.

[0012] There is further provided, in an embodiment, a computer program product comprising a non-transitory computer-readable storage medium having program instructions embodied therewith, the program instructions executable by at least one hardware processor to: receive a plurality of gaze vectors associated with a cohort of subjects, wherein each of the gaze vectors represents tracking of a point-of-gaze of one of the subjects while performing a cognitive task; and at a training stage, train a machine learning model on a training dataset comprising: (i) all of the gaze vectors, and (ii) annotations indicating a depression score of each of the subjects, to obtain a trained machine learning model configured to issue a prediction of depression in a target subject, based on target gaze vectors obtained from gaze tracking recording of the target subject while performing the cognitive task.

[0013] In some embodiments, the program instructions are further executable to apply, and the method further comprises applying, at an inference stage, the trained machine learning model to the target gaze vectors obtained from gaze tracking recording of the target subject while performing the cognitive task, to issue a prediction of depression in the target subject.

[0014] In some embodiments, the cohort of subjects comprises (i) a first subgroup of subjects having each depression symptoms, and (ii) a second subgroup of subjects having no depression symptoms.

[0015] In some embodiments, each of the gaze vectors represents a time-series of spatial coordinates of the point-of-gaze.

[0016] In some embodiments, each of the cognitive tasks comprises reading a sentence shown on a display, wherein the gaze vector represents a time-series of spatial coordinates of points at which each of the subjects focused its gaze relative to the sentence on the display during the reading.

[0017] In some embodiments, the each of the sentences represents a specific congruency-emotional condition, based on a chosen combination of emotional valence and sentence congruency.

[0018] In some embodiments, the specific congruency-emotional condition is selected from the group consisting of: congruent and positive, congruent and negative, incongruent and positive, incongruent and negative.

[0019] In some embodiments, each of the gaze vectors in the training dataset is further annotated with the congruency-emotional condition.

[0020] In some embodiments, the machine learning model comprises four congruency-emotional condition-specific sub-models, each trained on a training dataset comprising: (i) all of the gaze vectors associated with a respective one of the congruency-emotional conditions, and (ii) annotations indicating a depression score of the subject associated with each of the gaze vectors, to obtain four trained sub-models, each configured to issue a congruency-emotional condition-specific prediction of depression in a target subject, based on target gaze vectors obtained from gaze tracking recording of the target subject while performing the cognitive task associated with the respective one of the congruency-emotional conditions.

[0021] In some embodiments, with respect to each of the subjects, all of the gaze vectors associated with a respective one of the congruency-emotional conditions are aggregated into a single gaze vector.

[0022] In some embodiments, the program instructions are further executable to issue, and the method further comprises issuing, a combined prediction, based on aggregating all of the congruency-emotional condition-specific predictions by the four sub-models.

[0023] In some embodiments, the depression score with respect to each of the subjects represents one of: a PHQ-9 score of the subject, a binary score indicating the presence or absence of depression symptoms in the subject, or a binary score indicating depression symptoms in the subject above or below a median value.

[0024] In addition to the exemplary aspects and embodiments described above, further aspects and embodiments will become apparent by reference to the figures and by study of the following detailed description.

BRIEF DESCRIPTION OF THE FIGURES

[0025] The present invention will be understood and appreciated more comprehensively from the following detailed description taken in conjunction with the appended drawings in which:

[0026] FIG. 1 is a block diagram of an exemplary system for automated detection of depression tendency in a subject, based on tracking eye movement and/or a point-of-gaze of the subject while performing a cognitive (e.g., reading) task, in accordance with some embodiments of the present invention;

[0027] FIG. 2 is a flowchart which illustrates the functional steps in a method for automated detection of depression tendency in a subject, based on tracking eye movement and/or a point-of-gaze of the subject while performing a cognitive (e.g., reading) task, in accordance with some embodiments of the present invention;

[0028] FIG. 3 illustrates an exemplary network architecture which may be utilized by the present technique for automated detection of depression tendency in a subject, based on tracking eye movement and/or a point-of-gaze of the subject while performing a cognitive (e.g., reading) task, in accordance with some embodiments of the present invention;

[0029] FIGS. 4A-4D illustrate experimental results.

DETAILED DESCRIPTION

[0030] Disclosed herein is a technique, embodied as a system, method, and computer program product, for automated detection of depression tendency in a subject, based on tracking eye movement and/or a point-of-gaze of the subject while performing a cognitive (e.g., reading) task.

[0031] In some embodiments, the present technique is based on the insight that specific measures of attentional bias may be a cognitive factor that is abnormal in persons exhibiting depression tendencies, and therefore may be utilized to detect depression in subjects. Specifically, altered attentional biases in subjects with a tendency to depression may be measured using gaze tracking data collected during reading of sentences. In this regard, gaze tracking directly reflects a subject's attention, with a possible relation to the reward system, which was repeatedly shown to be deficient in subjects with depression. Thus, based on findings that tendency to depression influences allocation of attention in

various cognitive tasks, and on the assumption that reading is a serial process that heavily relies on attention, it may be deduced that tendency for depression will modulate reading patterns, at least in sentences specifically created to highlight attentional biases. Furthermore, the continuous nature of reading makes gaze tracking data particularly suitable for revealing these differences.

[0032] In some embodiments, the present technique provides for a trained machine learning model, configured for automated detection of a depression tendency in a subject, based on gaze tracking of the subject while performing a specified cognitive (e.g., reading) task.

[0033] In some embodiments, a machine learning model of the present technique may be trained, validated and tested on a training dataset comprising time-dependent signals associated with gaze tracking recorded with respect to a cohort of subjects while performing a specified cognitive (e.g., reading) task.

[0034] In some embodiments, the training dataset of the present technique is based on entire sequences of gaze tracking signals, rather than only on extracted features pertaining to aggregated measures associated with attention to defined regions-of-interest (ROI) within the sentences comprising the cognitive (e.g., reading) task.

[0035] Accordingly, in some embodiments, the present technique provides for a training dataset which comprises complete time-series of sequences of gaze tracking, rather than relying on self-extracted or self-engineered features. In this regard, the common approach to gaze tracking analysis during cognitive (e.g., reading) tasks includes decomposing the gaze tracking data into a set of discrete features that are specifically and manually defined based on the presented sentences, and consequently, require explicitly defining ROIs for each sentence. In contrast, the present technique utilizes the entire time-series rather than rely on feature-extraction techniques, and thus is able to learn latent patterns and exploit them for prediction, thereby improving assumed generalizability to the real world by increasing sensitivity. In addition, the present technique represents a more accessible way to analyze gaze tracking data, as it avoids the need to explicitly define the screen locations of ROIs and to calculate each specific feature, analysis steps which are prone to human errors and highly dependent on device calibration accuracy. Because the present technique classifies the data based on the temporal patterns in the signal, rather than on pixel-level ROI and discrete features,

the present machine learning method demands much lower technical capabilities and device requirements.

[0036] In some embodiments, the cohort of subjects represents cases with depression or depression tendency, as well as cases with no detected depression tendency. In some embodiments, a depression tendency in each particular subject may be validated based, e.g., on a standard questionnaire or test, such as the PHQ-9 (Patient Health Questionnaire-9). In some embodiments, the signals associated with gaze tracking comprising the training dataset, are labeled with labels indicating a depression tendency or lack thereof in each particular subject.

[0037] In some embodiments, the gaze tracking signals comprising the training dataset are recorded during a cognitive (e.g., reading) task performed by each subject in the cohort of subjects. In some embodiments, the cognitive (e.g., reading) task comprises sentences constructed to highlight expected attention biases in subjects with depression tendencies. Specifically, the cognitive (e.g., reading) task may comprise sentences with high emotional information, which may cause subjects with depression tendency to exhibit altered attentional bias when processing the information. In addition, disconfirmed predictions during sentence processing may interact with the emotional valence of a word, as reflected in processing difficulty resulting from re-analysis by the reader. Thus, the cognitive (e.g., reading) task may comprise sentences having emotionally-laden words (representing either a positive or a negative valence) in specific positions in the sentence, which either fulfill or contradict a previous expectation of the reader, to potentially induce altered attention in the subject.

[0038] In some embodiments, the present technique provides for a cognitive (e.g., reading) task comprising four sentence types, based on manipulating valence and semantic congruence in the sentence. For example, each sentence may have a positive or negative valence, as well as be semantically-congruent or incongruent, i.e., based on the harmony or agreement, or lack thereof, between the meaning of the sentence and the predictive expectation of the typical reader as derived from the context of the sentence.

[0039] In some embodiments, these sentence structures may be configured to elicit a valence-congruence interaction in subjects, which is typically moderated by attention, and thus affected by attentional bias and, as a direct consequence, by the level of depression. In other words, it is expected that different levels of tendency for depression

should lead to different allocation of gaze during reading of sentences based on these particular structures.

[0040] In this regard, a cognitive (e.g., reading) task may be more sensitive to the tendency for depression than other similar tasks, such as image viewing. Sentence reading relies on sequential information accumulation. Moreover, in sentence reading, the difficulty of processing at each time point can be easily altered to trigger various underlying abnormal mechanisms. In contrast, image processing is more encapsulated and lacks the detectable temporal aspects and dependencies which exist in sentences, and therefore cannot be as easily manipulated within stimuli. When using images, a single image is typically all positive or all negative, whereas a reading task may include a range of valences between positive and negative, which may be weighted-in during processing. Thus, using sentences instead of images might be a step towards creating a more sensitive measure to detect attentional biases.

[0041] A potential advantage of the present technique is, therefore, in that it provides an objective tool to measure attentional bias in a subject, which may be used to identify a tendency to depression in the subject. The present technique is based on using whole sequences of gaze tracking signals associated with a cognitive (e.g., reading) task, rather than features and measures pertaining only to specific ROIs in the cognitive (e.g., reading) task. Thus, the present technique provides a predictive tool that is more accurate and robust, based on using richer data with higher dimensionality.

[0042] In some embodiments, the present technique provides for better stimuli than commonly-used images for classification of depression tendency. As further discussed hereinbelow, using sentence stimuli increases sensitivity by providing a continuous, less encapsulated gaze sequence, which can then be used instead of manually extracted features. This procedure also decreases technological demands and improves robustness.

[0043] The present technique may be applied in addition, or as an alternative, to subjective diagnosis tools. In addition, the present technique may be applied as an objective validation measure, to estimate efficiency of treatments using cognitive training methods. In clinical practice, the present technique may be used as part of an annual check-up, as a decision support system for a psychiatrist, or post-diagnosis, to provide personalized treatment with on-going assessment. In some forms, the present

technique may be used as a self-administered diagnostic tool by user, for example, as an application executed on a mobile device.

[0044] FIG. 1 is a block diagram of an exemplary system **100** for training a machine learning model configured for automated detection of a depression tendency in a subject, based on tracking a point-of-gaze of the subject while performing a specified cognitive (e.g., reading) task, according to some embodiments of the present disclosure.

[0045] In some embodiments, system **100** may comprise a hardware processor **102**, and a random-access memory (RAM) **104**, and/or one or more non-transitory computer-readable storage device **106**. In some embodiments, system **100** may store in storage device **106** software instructions or components configured to operate a processing unit (also ‘hardware processor,’ ‘CPU,’ ‘quantum computer processor,’ or simply ‘processor’), such as hardware processor **102**. In some embodiments, the software components may include an operating system, including various software components and/or drivers for controlling and managing general system tasks (e.g., memory management, storage device control, power management, etc.) and facilitating communication between various hardware and software components.

[0046] The software instructions and/or components operating hardware processor **102** may comprise a language processing module **106a**, a gaze signal analysis module **106b**, and/or a machine learning module **106c**.

[0047] Language processing module **106a** may be configured to perform any desired language processing task, such as generating sentences intended for cognitive cognitive (e.g., reading) tasks, parsing sentences and determine underlying semantic meaning, performing textual analyses, and generating time-series vector representations of natural language.

[0048] Gaze signal analysis module **106b** may be configured to receive recorded gaze tracking signals associated with one or more subjects, and to process and analyze the received data. In some embodiments, the received gaze tracking data represents a point-of-gaze (i.e., the point at which a subject is looking), and/or the motion of a subject’s eye relative to the head. In some embodiments, the received gaze tracking data represents gaze tracking of one or more subjects while performing a cognitive task, such as a cognitive (e.g., reading) task. In some embodiments, the recorded gaze data may be in the form of a time-series indicating where the subject focused its gaze in each sampling

instance. In some embodiments, the recorded gaze tracking data may be acquired at a sampling frequency of between 100-2000 Hz, e.g., 1000 Hz.

[0049] Machine learning module **106c** may comprise any one or more neural networks (i.e., which include one or more neural network layers), and can be implemented to embody any appropriate neural network architecture and any suitable machine learning algorithm. In some embodiments, Machine learning module **106c** may be used to train, test, validate, and/or inference one or more machine learning model of the present technique.

[0050] In some embodiments, system **100** may further comprise a user interface **108** comprising, e.g., a display monitor for displaying images, a control panel for controlling system **100**, and a speaker for providing audio feedback.

[0051] System **100** as described herein is only an exemplary embodiment of the present invention, and in practice may be implemented in hardware only, software only, or a combination of both hardware and software. System **100** may have more or fewer components and modules than shown, may combine two or more of the components, or may have a different configuration or arrangement of the components. System **100** may include any additional component enabling it to function as an operable computer system, such as a motherboard, data busses, power supply, a network interface card, a display, an input device (e.g., keyboard, pointing device, touch-sensitive display), etc. (not shown). Components of system **100** may be co-located or distributed, or the system may be configured to run as one or more cloud computing ‘instances,’ ‘containers,’ ‘virtual machines,’ or other types of encapsulated software applications, as known in the art.

[0052] The instructions of system **100** will now be discussed with reference to the flowchart of FIG. 2 which illustrates the functional steps in a method **200** for training a machine learning model configured for automated detection of a depression tendency in a subject, based on tracking a point-of-gaze of the subject while performing a specified cognitive (e.g., reading) task, according to some embodiments of the present disclosure.

[0053] The various steps of method **200** will be described with continuous reference to exemplary system **100** shown in FIG. 1. The various steps of method **200** may either be performed in the order they are presented or in a different order (or even in parallel), as long as the order allows for a necessary input to a certain step to be obtained from an output of an earlier step. In addition, the steps of method **200** may be performed

automatically (e.g., by system **100** of FIG. 1), unless specifically stated otherwise. In addition, the steps of method **200** are set forth for exemplary purposes, and it is expected that modification to the flow chart is normally required to accommodate various network configurations and network carrier business policies.

[0054] Method **200** begins in step **202**, wherein the instructions of gaze signal analysis module **106b** may cause system **100** to receive, as input, a dataset comprising a plurality of eye/gaze tracking recordings associated with a cohort of subjects. In some embodiments, the cohort of subjects includes at least (i) a first subgroup representing cases exhibiting depression tendencies and/or symptoms, and (ii) a second subgroup representing cases not exhibiting depression tendencies and/or symptoms.

[0055] In some embodiments, the dataset of eye/gaze tracking recordings represent, with respect to each subject in the cohort, tracking of (i) a point-of-gaze, i.e., the point at which a subject is looking on a display or screen, and/or (ii) the motion of a subject's eye relative to the head, from which the point-of-gaze may be deduced. In each case, the tracking data is collected while the subject performs a cognitive task, such as a reading task, on a display monitor or a screen. In some embodiments, each recording of eye/gaze tracking may be in the form of a time-series indicating spatial coordinates (x, y) of the point or points at which the subject focused its gaze at each sampling instance over the duration of the recording. In some embodiments, the recorded eye/gaze tracking data may be acquired at a sampling frequency of between 100-2000 times per second, i.e., a sampling frequency or rate of 100-2000 Hz, during the performance of the cognitive task. In one example, the recorded eye/gaze tracking data may be acquired at a sampling frequency of 1,000 times per second, i.e., 1,000 Hz, during the performance of the cognitive task.

[0056] In some embodiments, the dataset of eye/gaze tracking recordings may be collected using any suitable system or device. For example, the eye/gaze tracking recordings may be collected using a video-based eye or gaze tracker. Video-based trackers comprise a video camera which focuses on one or both eyes and records eye movement as the viewer looks at a stimulus. Typically, video-based systems use the center of the pupil and infrared/near-infrared non-collimated light to create corneal reflections (CR). The vector between the pupil center and the corneal reflections can be used to compute the point of regard on surface or the gaze direction. A simple calibration procedure of the individual is usually needed before using the eye tracker. The camera

images are processed by software (as may be implemented by gaze signal analysis module **106b**) to generate gaze data.

[0057] In some embodiments, the input dataset of eye/gaze tracking recordings includes at least (i) a first subset of eye/gaze tracking recordings associated with each subject in the first subgroup of the cohort of subjects, which represents cases exhibiting depression tendencies and/or symptoms, and (ii) a second subset of eye/gaze tracking recordings associated with each subject in the second subgroup of the cohort of subjects, which represents cases not exhibiting depression tendencies and/or symptoms. In some embodiments, a depression tendency in each particular subject in the cohort may be validated based, e.g., on a standard questionnaire or test, such as the PHQ-9 (Patient Health Questionnaire-9). The PHQ-9 is a nine-item depression module derived from the full "Patient Health Questionnaire." Each of the items is scored on a 4-point Likert scale. PHQ-9 scores range from 0 to 27, with higher scores indicating greater severity of depression. Values of ≤ 4 , ≥ 5 , ≥ 10 and ≥ 15 indicate minimal, mild, moderate and severe depressive symptoms, respectively.

[0058] In some embodiments, the cognitive task performed during the eye/gaze tracking recording comprises a reading task over a plurality of sentences configured to test altered attentional biases in subjects. In some embodiments, the plurality of sentences comprising the cognitive (e.g., reading) task may be generated by system **100** based on the instructions of language processing module **106a**. In some embodiments, the generated cognitive tasks comprise a series of sentences configured to test altered attentional biases in subjects, wherein the plurality of sentences are generated according to a predefined format.

[0059] In one exemplary implementation, the instructions of language processing module **106a** may cause system **100** to generate a plurality of sentences in different congruency categories, each having different congruency-emotion conditionals associated therewith, i.e., various combinations of semantic congruency (e.g., whether the meaning of the sentence comports with the predictive expectation of the reader as derived from the context of the sentence) and emotional valence (e.g., positive or negative). In some embodiments, the various congruency categories may be configured to elicit a valence-congruence interaction in subjects, which is typically moderated by attention, and thus affected by attentional bias, which in turn may be affected by the level of depression in the subject. In other words, it is expected that different levels of tendency

for depression should lead to different allocation of gaze during reading of sentences based on these particular structures.

[0060] For example, the instructions of language processing module **106a** may cause system **100** to generate a plurality of sentences within the following four main sentence congruency categories, based on combinations of congruency-emotion conditionals:

- Congruency Category A: Congruent sentences with positive emotional valence.
- Congruency Category B: Congruent sentences with negative emotional valence.
- Congruency Category C: Incongruent sentences with positive emotional valence.
- Congruency Category D: Incongruent sentences with negative emotional valence.

[0061] Table 1 below shows several examples of sentences in the four main congruency categories, having various combinations of emotional valence and congruency.

Table 1:

Congruency Category	Congruency Condition	Sentiment of Target Word	Sentence Example
A	Congruent	Positive	The pool water was clean, and the kids enjoyed swimming in it
B		Negative	The pool water was dirty, and the kids were repelled by swimming in it
C	Incongruent	Positive	The pool water was dirty, and the kids enjoyed swimming in it
D		Negative	The pool water was clean, and the kids were repelled by swimming in it

[0062] In one exemplary instance, as will be further described below under “Experimental Results,” the present inventors obtained a plurality of eye/gaze tracking recordings from a cohort of 101 subjects, comprising 42 males and 59 females, with a mean age of 25.74 years. The cognitive (e.g., reading) task performed by each subject was selected from 32 sentence sets, each comprising four sentences according to the four

main congruency categories A-D described above, in which the valence and congruence were manipulated as detailed hereinabove.

[0063] All sentences were constructed such that a specific position in the beginning of the sentence (denoted as the source word) had an interaction with a subsequent word (denoted as the target word). The source word held either a positive or a negative valence and established a context for the target word. The target word likewise had either a positive or a negative valence, and was either congruent or incongruent based on the preceding context (see example in Table 1 above). It was expected that the misalignment in the incongruent conditions would cause a processing difficulty configured to elicit differences in processing strategy or allocation of attention by the subjects. The syntactic structure of the sentences within each set was identical, and the frequencies of the source and target words were controlled across the various sentence conditional iterations. Moreover, all sentences and all sets had approximately the same width, structure, and position of source and target words.

[0064] The cognitive (e.g., reading) task was structured such that each subject was presented with one sentence from each set for a total of 32 sentences, with the same number of sentences (eight) structured according to each one of the four sentence congruency categories A-D (as detailed in Table 1 above). To prevent strategic processing and priming, the cognitive (e.g., reading) task further included filler sentences representing natural, probable sentences. To ensure that subjects paid attention and were engaged in reading, they were asked simple yes/no comprehension questions following each sentence. Each subject completed the cognitive (e.g., reading) task at its own pace. At the beginning of each trial, a fixation cross was presented in the middle of the screen, followed by the sentence, centered and aligned around the middle of the screen. Subjects were not given instructions on what strategy to use to complete the task, except to press the space bar when they finish reading the sentence. The comprehension question was presented after the sentence, and subjects had to select the correct answer.

[0065] Accordingly, at the conclusion of step **202** system **100** may receive and store a dataset comprising a plurality of eye/gaze tracking recordings associated with the cohort of subjects. In some embodiments, each one of the eye/gaze tracking recordings may be associated and/or labeled with the following information:

- A particular subject of the cohort of subjects.

- A depression score depression score associated with the particular subject, indicating the presence of depression in the subject, which may be a PHQ-9 scores on a range from 0 to 27.
- A particular one of the four sentence congruency categories A-D (as detailed in Table 1 above).

[0066] With reference back to FIG. 2, in step **204**, the instructions of gaze signal analysis module **106b** may cause system **100** to perform a data preprocessing stage with respect to the eye/gaze tracking recordings included in the input dataset received in step **202**. In some embodiments, preprocessing comprises at least one of data cleaning and normalizing, removal of missing data, data quality control, data augmentations, and/or any other suitable preprocessing method or technique.

[0067] In one exemplary implementation, the instructions of gaze signal analysis module **106b** may cause system **100** to perform at least one of the following preprocessing steps, to normalize the eye/gaze tracking recordings in the dataset:

- Extracting a gaze time-series vector, with respect to each sentence in the cognitive (e.g., reading) task, representing the time-series of spatial coordinates (e.g., a horizontal sentence part) at which the subject focused its gaze at each sampling instance during the eye/gaze tracking recording associated with each sentence. In some embodiments, the time-series of spatial coordinates may comprise only x coordinates (e.g., along the horizontal dimension relative to a single-line sentence).
- Truncating or padding each time-series vector to a width of 3,500 datapoints, as needed.
- Normalizing the coordinate location to transform all x coordinates into the range $[0,1]$, representing a relative location within the sentence. This may be performed using the following equation:

$$x_{scaled} = \frac{x - \min(x)}{\max(x) - \min(x)}$$

- Smoothing the time-series vectors by averaging each set of four non-overlapping datapoints, to obtain a new value, thus decreasing the length of each time-series vector by a factor of 4, to 875 datapoints.

- Truncating the first 200 datapoints in each time-series vectors, to receive a final time-series vector of 675 datapoints. This is based on the assumption that the start of the eye/gaze tracking signal does not exhibit distinctive patterns across the cohort of subjects, and only represents the shift of the gaze to the beginning of the new sentence.

[0068] In some embodiments, step **204** may comprise aggregating, with respect to each subject in the cohort of subjects, all the time-series associated with each of the four congruency categories of sentences A-D (as listed in Table 1 above). Thus, for each of the subjects, there are created 4 congruency condition-specific time-series vectors, each representing the spatial x coordinates (e.g., a horizontal sentence part) at which the subject focused its gaze at each sampling instance during all the eye/gaze tracking recordings associated with a particular one of the four congruency categories of sentences A-D (as listed in Table 1 above).

[0069] At the conclusion of step **204**, the instructions of gaze signal analysis module **106b** may cause system **100** to obtain sets of one or more of the following:

- (i) Sentence-level time-series vectors associated with each of the subjects in the cohort, wherein each vector represents the time-series of spatial x coordinates (e.g., a horizontal sentence part) at which the subject focused its gaze at each sampling instance during the eye/gaze tracking recording associated with one sentence. In some embodiments, the set of sentence-level time-series vectors comprises 32 time-series vectors with respect to each subject in the cohort, representing eye/gaze tracking during a cognitive (e.g., reading) task comprising 32 sentences, divided into groups of 8 sentences associated with each of the congruency categories A-D listed in Table 1 above. In some embodiments, each one of the sentence-level time-series vectors may be associated and/or labeled with one or more of the following data points:
 - a. A particular subject of the cohort of subjects.
 - b. A depression score associated with the particular subject, indicating the presence of depression in the subject, which may be a PHQ-9 scores on a range from 0 to 27.
 - c. A particular one of the four sentence congruency categories A-D (as detailed in Table 1 above).

- (ii) congruency condition-specific time-series vectors associated with each of the subjects in the cohort, wherein each vector represents the aggregate time-series of spatial x coordinates (e.g., a horizontal sentence part) at which the subject focused its gaze at each sampling instance during all the eye/gaze tracking recordings associated with a particular one of the four congruency categories of sentences A-D (as listed in Table 1 above). In some embodiments, each one of the congruency condition-specific aggregate time-series vectors may be associated and/or labeled with one or more of the following data points:
 - a. A particular subject of the cohort of subjects.
 - b. A depression score associated with the particular subject, indicating the presence of depression in the subject, which may be a PHQ-9 scores on a range from 0 to 27.
 - c. A particular one of the four sentence congruency categories A-D (as detailed in Table 1 above)

[0070] With reference back to FIG. 2, in step **206**, the instructions of machine learning module **106c** may cause system **100** to construct one or more training datasets from the preprocessed time-series vectors obtained in step **204**.

[0071] Accordingly, in some embodiments, the instructions of machine learning module **106c** may cause system **100** to construct a first exemplary training dataset comprising:

- (i) The plurality of sentence-level time-series vectors with respect to all sentence congruency categories obtained in step **204**, associated with each of the subjects in the cohort; and
- (ii) labels indicating sentence congruency category (from the four congruency categories A-D listed in Table 1 above) and a depression score of the subject associated with each of the time-series vectors.

[0072] In some embodiments, the first exemplary training dataset may be configured to train a single machine learning model to predict depression tendency in a target subject, based on input data comprising a plurality of time-series vectors obtained from eye/gaze tracking of the subject while performing reading tasks comprising all four congruency categories of sentences A-D (as listed in Table 1 above).

[0073] In some embodiments, the instructions of machine learning module **106c** may cause system **100** to construct a second exemplary training dataset, comprising four congruency condition-specific sub-training datasets, each comprising only time-series vectors associated with a particular one of the four sentence congruency categories A-D (detailed in Table 1). Each of the four sub-training datasets thus comprises:

- (i) The plurality of congruency condition-specific time-series vectors obtained in step **204**, associated with each of the subjects in the cohort, in a particular one of the four sentence congruency categories A-D; and
- (ii) labels indicating a depression score of the subject associated with each of the time-series vectors.

[0074] In some embodiments, the congruency condition-specific time-series vectors included in the four congruency condition-specific sub-training datasets may be sentence-level vectors representing a time-series of spatial x coordinates (e.g., a horizontal sentence part) at which the subject focused its gaze at each sampling instance during an eye/gaze tracking recording of a reading task associated with one sentence.

[0075] In some embodiments, the time-series vectors included in the four congruency condition-specific sub-training datasets may be aggregate time-series, representing a time-series of spatial x coordinates (e.g., a horizontal sentence part) at which the subject focused its gaze at each sampling instance during all the eye/gaze tracking recordings associated with a particular one of the four congruency categories of sentences A-D (as listed in Table 1 above).

[0076] Accordingly, in some embodiments, the four congruency condition-specific sub-training datasets may comprise the following:

- (i) Sub-Dataset I-Congruent-Positive: Comprising time-series vectors associated with sentences from congruency category A (as detailed in Table 1). In some embodiments, all relevant time-series vectors with respect to each particular subject may be aggregated (e.g., averaged).
- (ii) Sub-Dataset II-Congruent-Negative: Comprising time-series vectors associated with sentences from congruency category B (as detailed in Table 1). In some embodiments, all relevant time-series vectors with respect to each particular subject may be aggregated (e.g., averaged).

- (iii) Sub-Dataset III-Incongruent-Positive: Comprising time-series vectors associated with sentences from congruency category C (as detailed in Table 1). In some embodiments, all relevant time-series vectors with respect to each particular subject may be aggregated (e.g., averaged).
- (iv) Sub-Dataset IV-Incongruent-Negative: Comprising time-series vectors associated with sentences from congruency category A (as detailed in Table 1). In some embodiments, all relevant time-series vectors with respect to each particular subject may be aggregated (e.g., averaged).

[0077] In some embodiments, one or more of the exemplary training datasets constructed in step **206** may be divided into a training portion, a validation portion, and/or a testing portion.

[0078] In each of the first and second exemplary training datasets detailed hereinabove, the depression score may represent each subject's PHQ-9 score. However, in some embodiments, the depression score may be indicated as a binary value (e.g., 0/1, yes/no), where a value of '0' or 'no' indicates no depression symptoms in the subject, and a value of '1' or 'yes' indicates the presence of depression symptoms in the subject. In other embodiments, the depression score may be indicated as a binary value (e.g., 0/1, yes/no), where a value of '0' or 'no' indicates depression symptoms in the subject below a median value, and a value of '1' or 'yes' indicates the presence of depression symptoms in the subject above a median value. In yet other embodiments, non-binary annotation may be used, e.g., on a scale of 1-5 or any other suitable scale, indicating a degree of severity of depression in a subject, e.g., none, mild, moderate, severe, very severe.

[0079] In some embodiments, in step **208**, the instructions of machine learning module **106c** may cause system **100** to train one or more machine learning models on the training datasets constructed in step **208**.

[0080] In some embodiments, one or more machine learning models of the present technique may be based on an exemplary network architecture, such as exemplary LSTM-based network **300** in FIG. 3. Exemplary architecture **300** may comprise a series of modules. Long short-term memory (LSTM) is a recurrent neural network (RNN) architecture that has feedback connections and can be used not only to process single data points, but also temporally related sequences of data (such as speech or video). This property makes it suitable for time-series analysis. A common LSTM unit is composed

of a cell, an input gate, an output gate and a forget gate. The cell remembers values over time, while the three gates moderate the flow of information into and out of the cell. In some embodiments, exemplary architecture **300** uses an LSTM network where every LSTM cell received as input the datapoint itself (for example, an x coordinate of a time-series vector) and an embedding of a sentence condition, which outputs eight dimensional vectors.

[0081] In some embodiments, the one or more machine learning models trained on the training datasets constructed in step **208**, may be configured to predict depression in a target subject, based on input data representing eye/gaze tracking of the target subject while performing a cognitive (e.g., reading) task.

[0082] In some embodiments, a first exemplary machine learning model of the present technique may be trained on the first exemplary training dataset constructed in step **208**.

[0083] In some embodiments, a second exemplary machine learning model of the present technique may be trained on the second exemplary training dataset constructed in step **206**. In some embodiments, the second exemplary machine learning model may comprise an ensemble of four sub-models, each trained on a particular one of the four sub-training datasets of the second exemplary training dataset constructed in step **208**, wherein each of the four sub-datasets comprises only congruency condition-specific time-series vectors associated with a particular one of the four sentence congruency categories A-D (detailed in Table 1).

[0084] Accordingly, in some embodiments, the second exemplary machine learning model may comprise the following four sub-models:

- (v) Sub-Model I-Congruent-Positive: Trained on Sub-dataset I, as detailed in step **206** above.
- (vi) Sub-Model II-Congruent-Negative: Trained on Sub-dataset II, as detailed in step **206** above.
- (vii) Sub-Model III-Incongruent-Positive: Trained on Sub-dataset III, as detailed in step **206** above.
- (viii) Sub-Model IV-Incongruent-Negative: Trained on Sub-dataset IV, as detailed in step **206** above.

[0085] With reference back to FIG. 2, in step **210**, the instructions of machine learning module **106c** may cause system **100** to perform an inference stage, wherein one or more of the machine learning models of the present technique as trained in step **208** may be applied to target input data comprising one or more time-series vectors associated with a target subject, to predict the presence or absence of depression symptoms in the target subject

[0086] In some embodiments, target input data comprising a at least one time-series vector are obtained based on eye/gaze tracking recordings of the target subject while performing reading tasks representing sentences selected from the four congruency categories of A-D (as listed in Table 1 above).

[0087] In some embodiments, the obtained eye/gaze tracking recordings of the target subject may undergo a data preprocessing stage as detailed in step **204** hereinabove, to produce the target input data, comprising a set of:

- (i) Sentence-level time-series vectors associated with the target subject, wherein each vector represents the time-series of spatial x coordinates (e.g., a horizontal sentence part) at which the target subject focused its gaze at each sampling instance during the eye/gaze tracking recording associated with one sentence. In some embodiments, the set of sentence-level time-series vectors is associated with sentences selected from each of the congruency categories A-D listed in Table 1 above; and/or
- (ii) congruency condition-specific time-series vectors associated with the target subject, wherein each vector represents the aggregate time-series of spatial x coordinates (e.g., a horizontal sentence part) at which the target subject focused its gaze at each sampling instance during all the eye/gaze tracking recordings associated with a particular one of the four congruency categories of sentences A-D (as listed in Table 1 above).

[0088] In some embodiments, in step **210**, the instructions of machine learning module **106c** may cause system **100** to apply the first exemplary machine learning model trained in step **208** to the target input data comprising sentence-level time-series vectors associated with the target subject, wherein each vector represents the time-series of spatial x coordinates (e.g., a horizontal sentence part) at which the target subject focused its gaze at each sampling instance during the eye/gaze tracking recording associated with

one sentence, and wherein the set of sentence-level time-series vectors is associated with sentences selected from each of the congruency categories A-D listed in Table 1 above.

[0089] In some embodiments, the prediction issued by the trained first exemplary machine learning model indicates the presence or absence of depression symptoms in the target subject. In some embodiments, the prediction issued by the trained first exemplary machine learning model indicates the presence or absence of depression symptoms in the target subject, as associated with a confidence score which represents the likelihood that the output of the machine learning model is correct.

[0090] In some embodiments, the trained machine learning model may be configured to issue a separate prediction with respect to each of the plurality of sentence-level time-series vectors comprising the target input data, wherein each of the separate predictions indicates the presence or absence of depression symptoms in the target subject. In some embodiments, the machine learning model is then configured to issue a final prediction representing an aggregation (for example, by averaging) of the multiple separate predictions, wherein the final prediction indicates the presence or absence of depression symptoms in the target subject.

[0091] In some embodiments, in the case where the labeling scheme of the training dataset used to train the first exemplary machine learning model uses actual PHQ-9 scores, the prediction issued by the trained machine learning model indicates the PHQ-9 score of the target subject.

[0092] In cases where the labeling scheme of the training dataset used to train the first exemplary machine learning model uses a binary value (e.g., 0/1, yes/no), where a value of '0' or 'no' indicates no depression symptoms in the subject, and a value of '1' or 'yes' indicates the presence of depression symptoms in the subject, the prediction issued by the trained machine learning model indicates the likelihood that the target subject has depression.

[0093] In cases where the labeling scheme of the training dataset used to train the first exemplary machine learning model uses a binary value (e.g., 0/1, yes/no), where a value of '0' or 'no' indicates depression symptoms in the subject below a median value, and a value of '1' or 'yes' indicates the presence of depression symptoms in the subject above a median value, the prediction issued by the trained machine learning model indicates the likelihood that the target subject has a PHQ-9 score above or below the median value.

[0094] In some embodiments, in step **210**, the instructions of machine learning module **106c** may cause system **100** to apply the second exemplary machine learning model trained in step **208**, comprising four sub-models I-IV, to the target input data comprising congruency condition-specific time-series vectors.

[0095] In some embodiments, the instructions of machine learning module **106c** may cause system **100** to apply the each of the four sub-models I-IV respectively to congruency condition-specific target time-series vectors associated with the congruency category of sentences A-D (as listed in Table 1 above) corresponding to the one on which each of the sub-models I-IV was trained.

[0096] In some embodiments, the time-series vectors may be aggregate time-series, representing a time-series of spatial x coordinates (e.g., a horizontal sentence part) at which the subject focused its gaze at each sampling instance during all the eye/gaze tracking recordings associated with a particular one of the four congruency categories of sentences A-D (as listed in Table 1 above).

[0097] In some embodiments, each of the trained sub-models I-IV comprising the second exemplary machine learning model may be configured to issue a prediction which indicates the presence or absence of depression symptoms in the target subject. In some embodiments, each of the predictions is associated with a confidence score which represents the likelihood that the output of the machine learning model is correct.

[0098] In some embodiments, the respective predictions issued by each of the trained sub-models I-IV comprising the second exemplary machine learning model may be aggregated into a single final prediction, e.g., by averaging the individual predictions.

[0099] In some embodiments, in the case where the labeling schemes of the training datasets used to train the respective sub-models comprising the second exemplary machine learning model use actual PHQ-9 scores, the respective predictions indicate the PHQ-9 score of the target subject.

[00100] In cases where the labeling schemes of the training datasets used to train the respective sub-models comprising the second exemplary machine learning model use a binary value (e.g., 0/1, yes/no), where a value of '0' or 'no' indicates no depression symptoms in the subject, and a value of '1' or 'yes' indicates the presence of depression symptoms in the subject, the predictions indicate the likelihood that the target subject has depression.

[00101] In cases where the labeling schemes of the training datasets used to train the respective sub-models comprising the second exemplary machine learning model use a binary value (e.g., 0/1, yes/no), where a value of '0' or 'no' indicates depression symptoms in the subject below a median value, and a value of '1' or 'yes' indicates the presence of depression symptoms in the subject above a median value, the predictions indicate the likelihood that the target subject has a PHQ-9 score above or below the median value.

Experimental Results

[00102] The present inventors conducted an experiment with respect to the exemplary cohort and cognitive (e.g., reading) task parameters detailed with reference to step **202** of method **200** hereinabove.

[00103] FIG. 4A show a histogram of the distribution of PHQ scores in the cohort of subjects.

[00104] The present inventors used two exemplary machine learning model algorithms, trained on the training datasets (as detailed in step **208** of method **200** hereinabove), as follows:

- Random Forest: Using a random forest classifier, prediction accuracy was 63% (std=0.09), with ROC-AUC of 0.66 with $p < 0.001$. This suggests it is possible to detect PHQ labels from reading gaze patterns). Moreover, as can be seen in FIG. 4D, per-subject accuracy is lower when the PHQ score is closer to the classification threshold (median). This is compared with 47% accuracy (std=0.089) for a random forest validation model trained on null distribution, generated by shuffling the labels used to annotate the data in the training datasets.
- LSTM with condition embeddings: Using LSTM-based models, prediction accuracy was 65% (std=0.14), with $p < 0.001$. Prediction accuracy was 63% (std=0.09), with ROC-AUC of 0.66 with $p < 0.001$. This is compared with 45% accuracy (std=0.13) for an LSTM validation model trained on null distribution, generated by shuffling the labels used to annotate the data in the training datasets.

[00105] FIGS. 4B-4C illustrate accuracy results.

[00106] FIG. 4B is a box plot of fold score between models trained on actual data (left hand side) and on null distribution (right hand side, generated by shuffling the labels used to annotate the data in the training datasets) for the Random Forest model.

[00107] FIG. 4C shows a box plot of fold score between model trained on actual data (left hand side) and on null distribution (right hand side, generated by shuffling the labels) for the LSTM model, with separate entries for congruent condition.

[00108] FIG. 4D illustrates the relation between subjects' PHQ scores and the accuracy of corresponding subjects. Bins that are closer to the median PHQ score (7, as shown in FIG. 4A) show more false classifications.

[00109] The present invention may be a system, a method, and/or a computer program product. The computer program product may include a computer readable storage medium (or media) having computer readable program instructions thereon for causing a processor to carry out aspects of the present invention.

[00110] The computer readable storage medium can be a tangible device that can retain and store instructions for use by an instruction execution device. The computer readable storage medium may be, for example, but is not limited to, an electronic storage device, a magnetic storage device, an optical storage device, an electromagnetic storage device, a semiconductor storage device, or any suitable combination of the foregoing. A non-exhaustive list of more specific examples of the computer readable storage medium includes the following: a portable computer diskette, a hard disk, a random access memory (RAM), a read-only memory (ROM), an erasable programmable read-only memory (EPROM or Flash memory), a static random access memory (SRAM), a portable compact disc read-only memory (CD-ROM), a digital versatile disk (DVD), a memory stick, a floppy disk, a mechanically encoded device having instructions recorded thereon, and any suitable combination of the foregoing. A computer readable storage medium, as used herein, is not to be construed as being transitory signals per se, such as radio waves or other freely propagating electromagnetic waves, electromagnetic waves propagating through a waveguide or other transmission media (e.g., light pulses passing through a fiber-optic cable), or electrical signals transmitted through a wire. Rather, the computer readable storage medium is a non-transient (i.e., not-volatile) medium.

[00111] Computer readable program instructions described herein can be downloaded to respective computing/processing devices from a computer readable storage medium

or to an external computer or external storage device via a network, for example, the Internet, a local area network, a wide area network and/or a wireless network. The network may comprise copper transmission cables, optical transmission fibers, wireless transmission, routers, firewalls, switches, gateway computers and/or edge servers. A network adapter card or network interface in each computing/processing device receives computer readable program instructions from the network and forwards the computer readable program instructions for storage in a computer readable storage medium within the respective computing/processing device.

[00112] Computer readable program instructions for carrying out operations of the present invention may be assembler instructions, instruction-set-architecture (ISA) instructions, machine instructions, machine dependent instructions, microcode, firmware instructions, state-setting data, or either source code or object code written in any combination of one or more programming languages, including an object-oriented programming language such as Java, Smalltalk, C++ or the like, and conventional procedural programming languages, such as the "C" programming language or similar programming languages. The computer readable program instructions may execute entirely on the user's computer, partly on the user's computer, as a stand-alone software package, partly on the user's computer and partly on a remote computer or entirely on the remote computer or server. In the latter scenario, the remote computer may be connected to the user's computer through any type of network, including a local area network (LAN) or a wide area network (WAN), or the connection may be made to an external computer (for example, through the Internet using an Internet Service Provider). In some embodiments, electronic circuitry including, for example, programmable logic circuitry, a field-programmable gate array (FPGA), or a programmable logic array (PLA) may execute the computer readable program instructions by utilizing state information of the computer readable program instructions to personalize the electronic circuitry, in order to perform aspects of the present invention. In some embodiments, electronic circuitry including, for example, an application-specific integrated circuit (ASIC), may be incorporate the computer readable program instructions already at time of fabrication, such that the ASIC is configured to execute these instructions without programming.

[00113] Aspects of the present invention are described herein with reference to flowchart illustrations and/or block diagrams of methods, apparatus (systems), and computer program products according to embodiments of the invention. It will be

understood that each block of the flowchart illustrations and/or block diagrams, and combinations of blocks in the flowchart illustrations and/or block diagrams, can be implemented by computer readable program instructions.

[00114] These computer readable program instructions may be provided to a processor of a general-purpose computer, special purpose computer, or other programmable data processing apparatus to produce a machine, such that the instructions, which execute via the processor of the computer or other programmable data processing apparatus, create means for implementing the functions/acts specified in the flowchart and/or block diagram block or blocks. These computer readable program instructions may also be stored in a computer readable storage medium that can direct a computer, a programmable data processing apparatus, and/or other devices to function in a particular manner, such that the computer readable storage medium having instructions stored therein comprises an article of manufacture including instructions which implement aspects of the function/act specified in the flowchart and/or block diagram block or blocks.

[00115] The computer readable program instructions may also be loaded onto a computer, other programmable data processing apparatus, or other device to cause a series of operational steps to be performed on the computer, other programmable apparatus or other device to produce a computer implemented process, such that the instructions which execute on the computer, other programmable apparatus, or other device implement the functions/acts specified in the flowchart and/or block diagram block or blocks.

[00116] The flowchart and block diagrams in the Figures illustrate the architecture, functionality, and operation of possible implementations of systems, methods, and computer program products according to various embodiments of the present invention. In this regard, each block in the flowchart or block diagrams may represent a module, segment, or portion of instructions, which comprises one or more executable instructions for implementing the specified logical function(s). It will also be noted that each block of the block diagrams and/or flowchart illustration, and combinations of blocks in the block diagrams and/or flowchart illustration, can be implemented by special purpose hardware-based systems that perform the specified functions or acts or carry out combinations of special purpose hardware and computer instructions.

[00117] In the description and claims, each of the terms “substantially,” “essentially,” and forms thereof, when describing a numerical value, means up to a 20% deviation (namely, $\pm 20\%$) from that value. Similarly, when such a term describes a numerical range, it means up to a 20% broader range – 10% over that explicit range and 10% below it).

[00118] In the description, any given numerical range should be considered to have specifically disclosed all the possible subranges as well as individual numerical values within that range, such that each such subrange and individual numerical value constitutes an embodiment of the invention. This applies regardless of the breadth of the range. For example, description of a range of integers from 1 to 6 should be considered to have specifically disclosed subranges such as from 1 to 3, from 1 to 4, from 1 to 5, from 2 to 4, from 2 to 6, from 3 to 6, etc., as well as individual numbers within that range, for example, 1, 4, and 6. Similarly, description of a range of fractions, for example from 0.6 to 1.1, should be considered to have specifically disclosed subranges such as from 0.6 to 0.9, from 0.7 to 1.1, from 0.9 to 1, from 0.8 to 0.9, from 0.6 to 1.1, from 1 to 1.1 etc., as well as individual numbers within that range, for example 0.7, 1, and 1.1.

[00119] The descriptions of the various embodiments of the present invention have been presented for purposes of illustration, but are not intended to be exhaustive or limited to the explicit descriptions. Many modifications and variations will be apparent to those of ordinary skill in the art without departing from the scope and spirit of the described embodiments. The terminology used herein was chosen to best explain the principles of the embodiments, the practical application or technical improvement over technologies found in the marketplace, or to enable others of ordinary skill in the art to understand the embodiments disclosed herein.

[00120] In the description and claims of the application, each of the words “comprise,” “include,” and “have,” as well as forms thereof, are not necessarily limited to members in a list with which the words may be associated.

[00121] Where there are inconsistencies between the description and any document incorporated by reference or otherwise relied upon, it is intended that the present description controls.

CLAIMS

What is claimed is:

1. A system comprising:
 - at least one hardware processor; and
 - a non-transitory computer-readable storage medium having stored thereon program instructions, the program instructions executable by the at least one hardware processor to:
 - receive a plurality of gaze vectors associated with a cohort of subjects, wherein each of said gaze vectors represents tracking of a point-of-gaze of one of said subjects while performing a cognitive task, and
 - at a training stage, train a machine learning model on a training dataset comprising:
 - (i) all of said gaze vectors, and
 - (ii) annotations indicating a depression score of each of said subjects,to obtain a trained machine learning model configured to issue a prediction of depression in a target subject, based on target gaze vectors obtained from gaze tracking recording of said target subject while performing said cognitive task.
2. The system of claim 1, wherein said program instructions are further executable to apply, at an inference stage, said trained machine learning model to said target gaze vectors obtained from gaze tracking recording of said target subject while performing said cognitive task, to issue a prediction of depression in said target subject.
3. The system of any one of claims 1 or 2, wherein said cohort of subjects comprises (i) a first subgroup of subjects having each depression symptoms, and (ii) a second subgroup of subjects having no depression symptoms.
4. The system of any one of claims 1-3, wherein each of said gaze vectors represents a time-series of spatial coordinates of said point-of-gaze.
5. The system of any one of claims 1-4, wherein each of said cognitive tasks comprises reading a sentence shown on a display, and wherein said gaze vector represents a time-series of spatial coordinates of points at which each of said subjects focused its gaze relative to said sentence on said display during said reading.

6. The system of claim 5, wherein each of said sentences represents a specific congruency-emotional condition, based on a chosen combination of emotional valence and sentence congruency.

7. The system of claim 6, wherein said specific congruency-emotional condition is selected from the group consisting of: congruent and positive, congruent and negative, incongruent and positive, incongruent and negative.

8. The system of claim 7, wherein each of said gaze vectors in said training dataset is further annotated with said congruency-emotional condition.

9. The system of any one of claims 7 or 8, wherein said machine learning model comprises four congruency-emotional condition-specific sub-models, each trained on a training dataset comprising:

- (i) all of said gaze vectors associated with a respective one of said congruency-emotional conditions, and
- (ii) annotations indicating a depression score of said subject associated with each of said gaze vectors,

to obtain four trained sub-models, each configured to issue a congruency-emotional condition-specific prediction of depression in a target subject, based on target gaze vectors obtained from gaze tracking recording of said target subject while performing said cognitive task associated with said respective one of said congruency-emotional conditions.

10. The system of claim 9, wherein, with respect to each of said subjects, all of said gaze vectors associated with a respective one of said congruency-emotional conditions are aggregated into a single gaze vector.

11. The system of any one of claims 9 or 10, wherein said program instructions are further executable to issue a combined prediction, based on aggregating all of said congruency-emotional condition-specific predictions by said four sub-models.

12. The system of any one of claims 1-11, wherein said depression score with respect to each of said subjects represents one of: a PHQ-9 score of said subject, a binary score indicating the presence or absence of depression symptoms in said subject, or a binary score indicating depression symptoms in the subject above or below a median value.

13. A computer-implemented method comprising:

receiving a plurality of gaze vectors associated with a cohort of subjects, wherein each of said gaze vectors represents tracking of a point-of-gaze of one of said subjects while performing a cognitive task; and

at a training stage, training a machine learning model on a training dataset comprising:

(i) all of said gaze vectors, and

(ii) annotations indicating a depression score of each of said subjects,

to obtain a trained machine learning model configured to issue a prediction of depression in a target subject, based on target gaze vectors obtained from gaze tracking recording of said target subject while performing said cognitive task.

14. The computer-implemented method of claim 13, further comprising applying, at an inference stage, said trained machine learning model to said target gaze vectors obtained from gaze tracking recording of said target subject while performing said cognitive task, to issue a prediction of depression in said target subject.

15. The computer-implemented method of any one of claims 13 or 14, wherein said cohort of subjects comprises (i) a first subgroup of subjects having each depression symptoms, and (ii) a second subgroup of subjects having no depression symptoms.

16. The computer-implemented method of any one of claims 13-15, wherein each of said gaze vectors represents a time-series of spatial coordinates of said point-of-gaze.

17. The computer-implemented method of any one of claims 13-16, wherein each of said cognitive tasks comprises reading a sentence shown on a display, and wherein said gaze vector represents a time-series of spatial coordinates of points at which each of said subjects focused its gaze relative to said sentence on said display during said reading.

18. The computer-implemented method of claim 17, wherein each of said sentences represents a specific congruency-emotional condition, based on a chosen combination of emotional valence and sentence congruency.

19. The computer-implemented method of claim 18, wherein said specific congruency-emotional condition is selected from the group consisting of: congruent and positive, congruent and negative, incongruent and positive, incongruent and negative.

20. The computer-implemented method of claim 19, wherein each of said gaze vectors in said training dataset is further annotated with said congruency-emotional condition.

21. The computer-implemented method of claim 19 or 20, wherein said machine learning model comprises four congruency-emotional condition-specific sub-models, each trained on a training dataset comprising:

- (i) all of said gaze vectors associated with a respective one of said congruency-emotional conditions, and
- (ii) annotations indicating a depression score of said subject associated with each of said gaze vectors,

to obtain four trained sub-models, each configured to issue a congruency-emotional condition-specific prediction of depression in a target subject, based on target gaze vectors obtained from gaze tracking recording of said target subject while performing said cognitive task associated with said respective one of said congruency-emotional conditions.

22. The computer-implemented method of claim 21, wherein, with respect to each of said subjects, all of said gaze vectors associated with a respective one of said congruency-emotional conditions are aggregated into a single gaze vector.

23. The computer-implemented method of any one of claims 21 or 22, further comprising issuing a combined prediction, based on aggregating all of said congruency-emotional condition-specific predictions by said four sub-models.

24. The computer-implemented method of any one of claims 13-23, wherein said depression score with respect to each of said subjects represents one of: a PHQ-9 score of said subject, a binary score indicating the presence or absence of depression symptoms in said subject, or a binary score indicating depression symptoms in the subject above or below a median value.

25. A computer program product comprising a non-transitory computer-readable storage medium having program instructions embodied therewith, the program instructions executable by at least one hardware processor to:

receive a plurality of gaze vectors associated with a cohort of subjects, wherein each of said gaze vectors represents tracking of a point-of-gaze of one of said subjects while performing a cognitive task; and

at a training stage, train a machine learning model on a training dataset comprising:

- (i) all of said gaze vectors, and
- (ii) annotations indicating a depression score of each of said subjects,

to obtain a trained machine learning model configured to issue a prediction of depression in a target subject, based on target gaze vectors obtained from gaze tracking recording of said target subject while performing said cognitive task.

26. The computer program product of claim 25, wherein said program instructions are further executable to apply, at an inference stage, said trained machine learning model to said target gaze vectors obtained from gaze tracking recording of said target subject while performing said cognitive task, to issue a prediction of depression in said target subject.

27. The computer program product of any one of claims 25 or 26, wherein said cohort of subjects comprises (i) a first subgroup of subjects having each depression symptoms, and (ii) a second subgroup of subjects having no depression symptoms.

28. The computer program product of any one of claims 25-27, wherein each of said gaze vectors represents a time-series of spatial coordinates of said point-of-gaze.

29. The computer program product of any one of claims 25-28, wherein each of said cognitive tasks comprises reading a sentence shown on a display, and wherein said gaze vector represents a time-series of spatial coordinates of points at which each of said subjects focused its gaze relative to said sentence on said display during said reading.

30. The computer program product of claim 29, wherein each of said sentences represents a specific congruency-emotional condition, based on a chosen combination of emotional valence and sentence congruency.

31. The computer program product of claim 30, wherein said specific congruency-emotional condition is selected from the group consisting of: congruent and positive, congruent and negative, incongruent and positive, incongruent and negative.

32. The computer program product of claim 31, wherein each of said gaze vectors in said training dataset is further annotated with said congruency-emotional condition.

33. The computer program product of claim 31 or 32, wherein said machine learning model comprises four congruency-emotional condition-specific sub-models, each trained on a training dataset comprising:

- (i) all of said gaze vectors associated with a respective one of said congruency-emotional conditions, and
- (ii) annotations indicating a depression score of said subject associated with each of said gaze vectors,

to obtain four trained sub-models, each configured to issue a congruency-emotional condition-specific prediction of depression in a target subject, based on target gaze vectors obtained from gaze tracking recording of said target subject while performing said cognitive task associated with said respective one of said congruency-emotional conditions.

34. The computer program product of claim 33, wherein, with respect to each of said subjects, all of said gaze vectors associated with a respective one of said congruency-emotional conditions are aggregated into a single gaze vector.

35. The computer program product of any one of claims 33 or 34, wherein said program instructions are further executable to issue a combined prediction, based on aggregating all of said congruency-emotional condition-specific predictions by said four sub-models.

36. The computer program product of any one of claims 25-35, wherein said depression score with respect to each of said subjects represents one of: a PHQ-9 score of said subject, a binary score indicating the presence or absence of depression symptoms in said subject, or a binary score indicating depression symptoms in the subject above or below a median value.

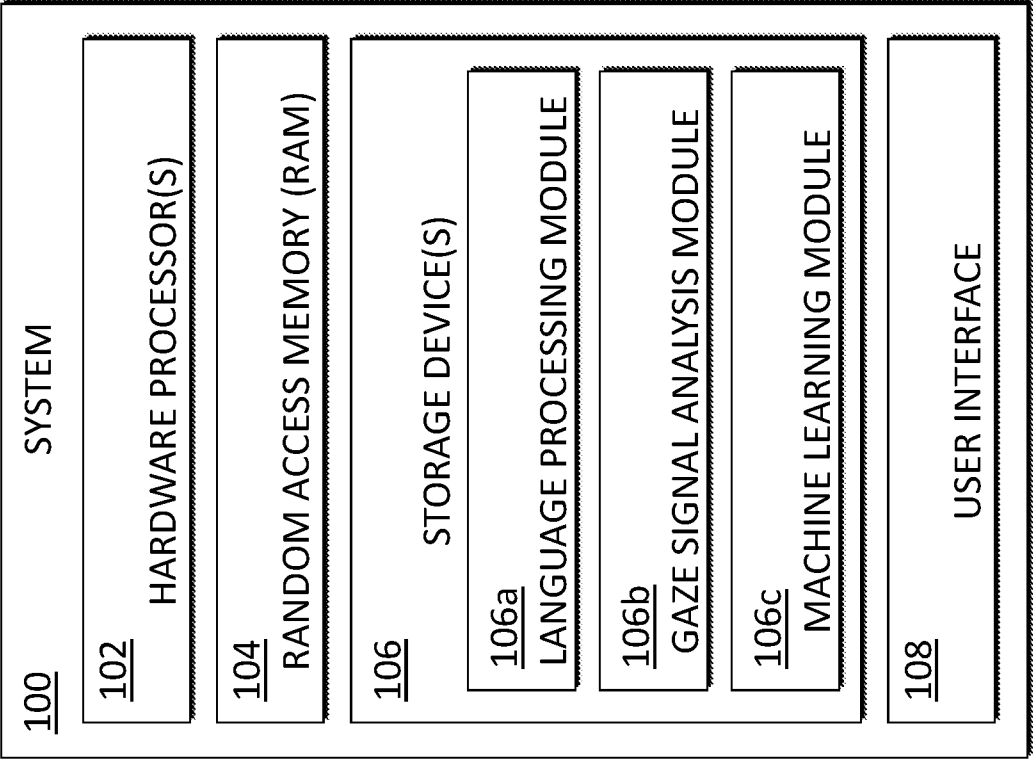


FIG. 1

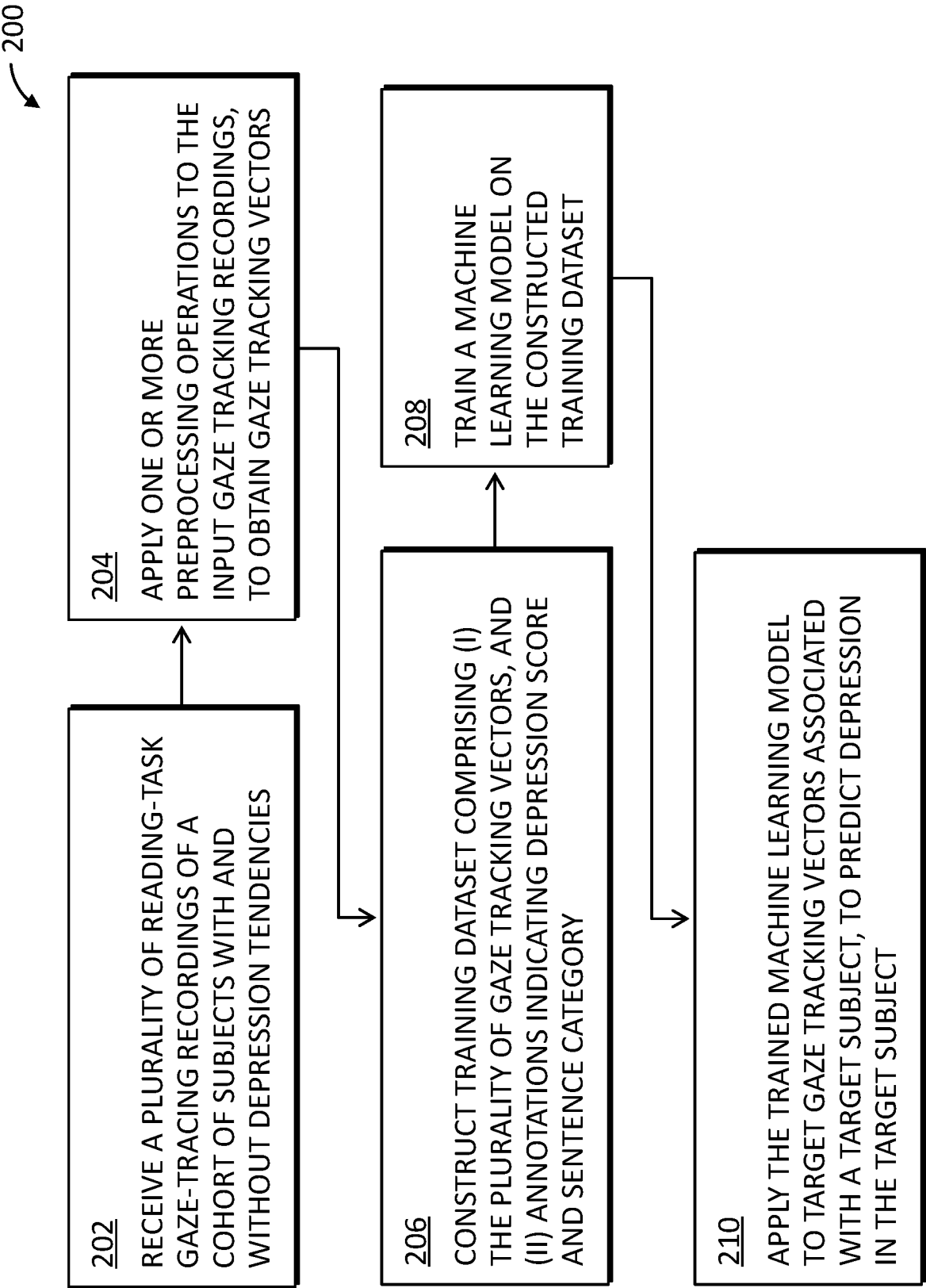


FIG. 2

300

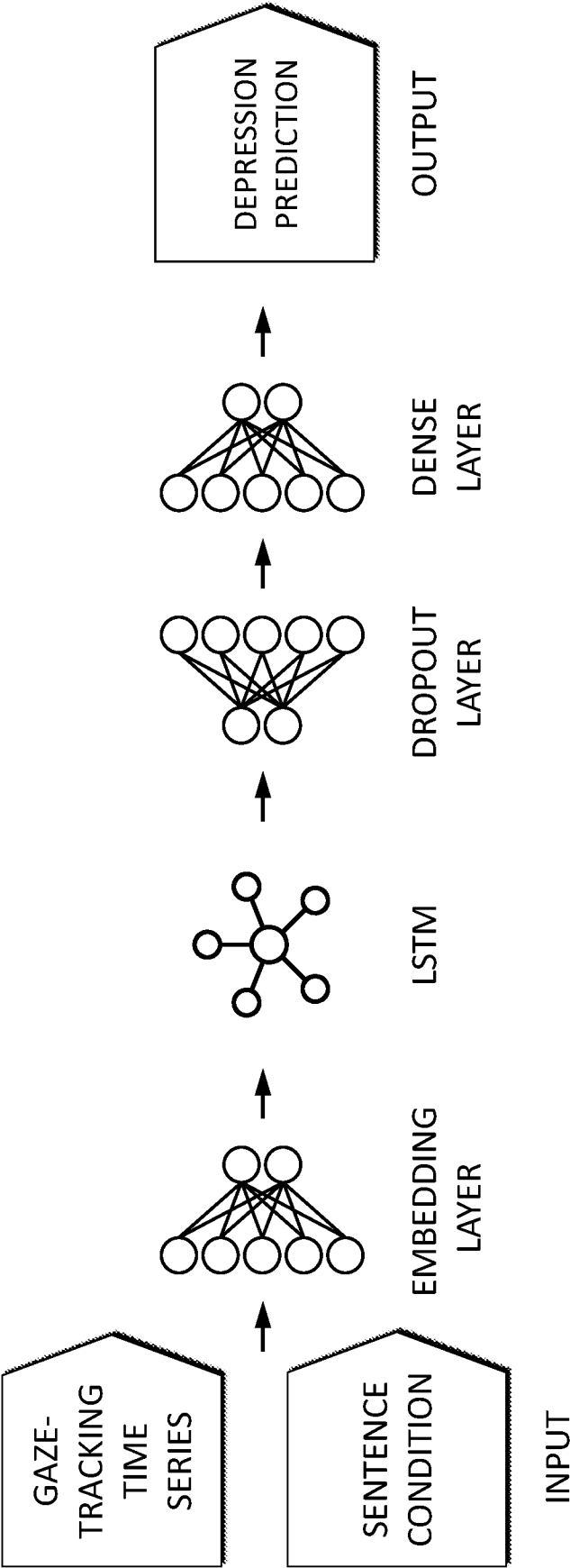


FIG. 3

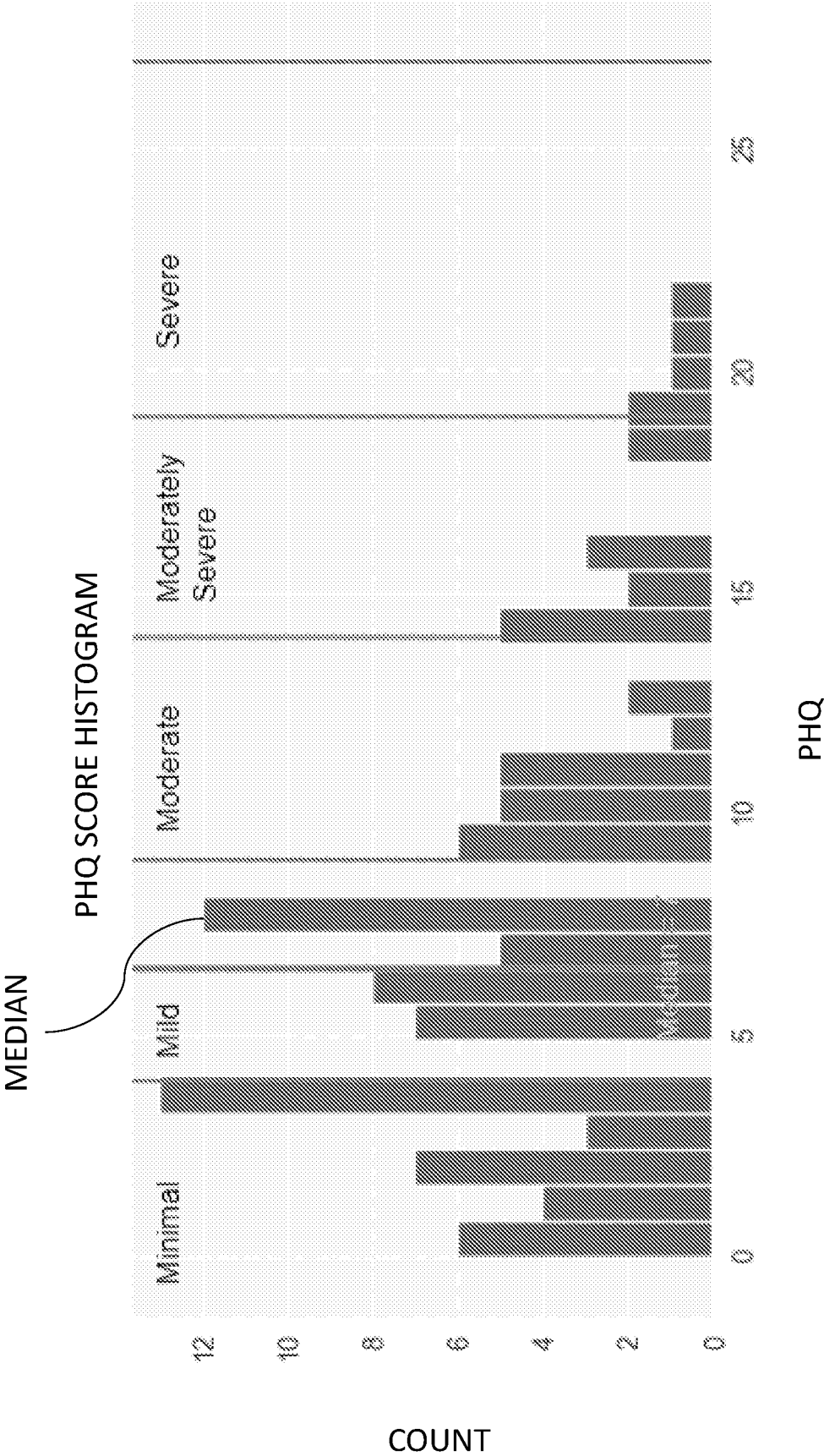


FIG. 4A

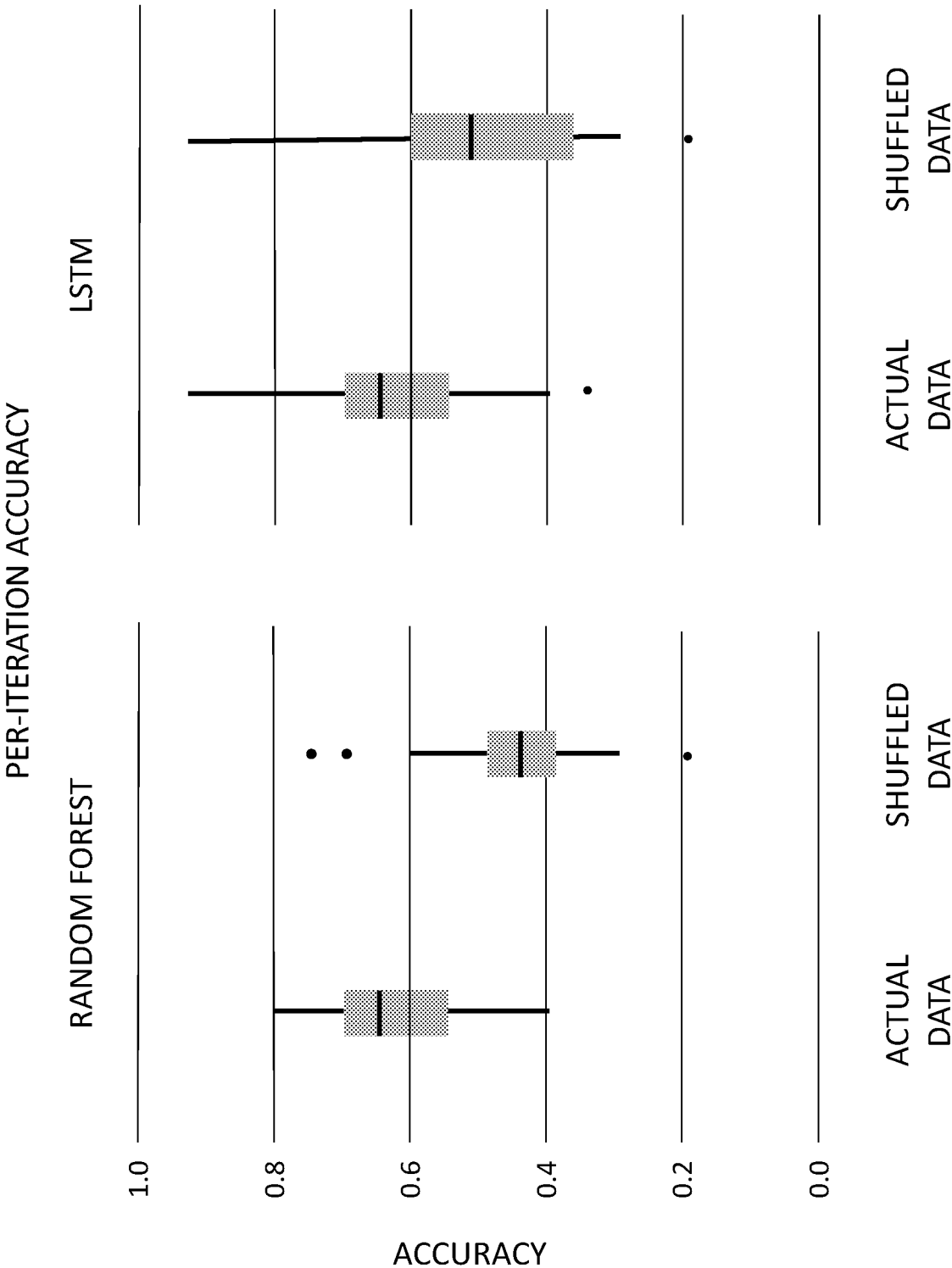


FIG. 4B

FIG. 4C

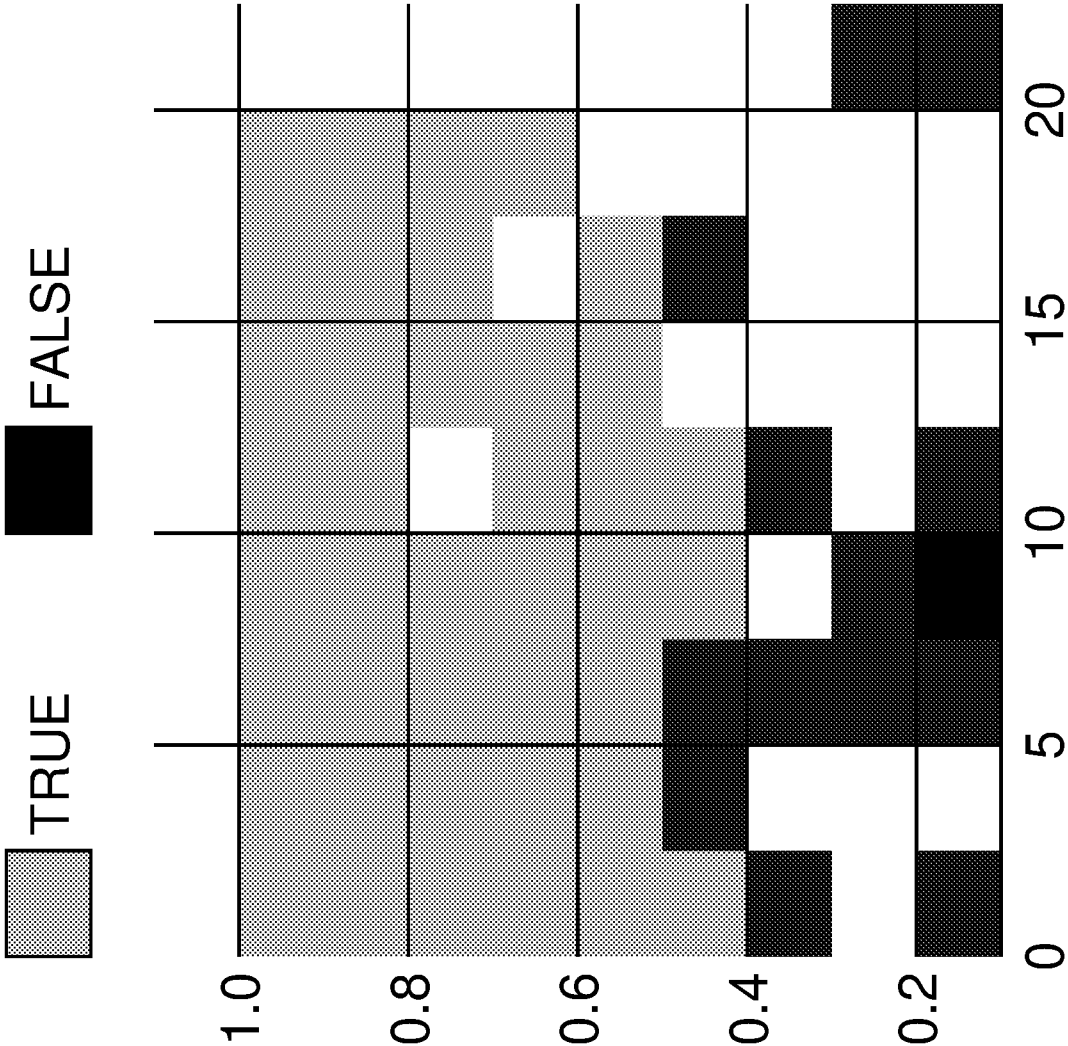


FIG. 4D

INTERNATIONAL SEARCH REPORT

International application No.

PCT/IL2024/050611

A. CLASSIFICATION OF SUBJECT MATTER A61B 5/00 (2024.01)i; A61B 5/16 (2024.01)i; G16H 40/67 (2024.01)i CPC:A61B 5/00; A61B 5/16; G16H 40/67 According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) A61B 5/00; G16H 40/00 CPC:A61B 5/00; G16H 40/00 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) Databases consulted: Esp@cenet, Google Patents, Orbit		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2021085180 A1 (WINTERLIGHT LABS INC [CA] et al) 25 March 2021 (2021-03-25) whole document	1-36
A	CA 3193776 A1 (LINUS HEALTH INC [US]) 31 March 2022 (2022-03-31) whole document	1-36
☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.		
* Special categories of cited documents: “A” document defining the general state of the art which is not considered to be of particular relevance “D” document cited by the applicant in the international application “E” earlier application or patent but published on or after the international filing date “L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) “O” document referring to an oral disclosure, use, exhibition or other means “P” document published prior to the international filing date but later than the priority date claimed “T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention “X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone “Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art “&” document member of the same patent family		
Date of the actual completion of the international search 08 October 2024		Date of mailing of the international search report 09 October 2024
Name and mailing address of the ISA/IL Israel Patent Office Technology Park, Bldg.5, Malcha, Jerusalem, 9695101, Israel Israel Telephone No. 972-73-3927253 Email: pctoffice@justice.gov.il		Authorized officer ZOZULYA Irina Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/IL2024/050611

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
US	2021085180	A1	25 March 2021	US	2021085180	A1	25 March 2021
				EP	1975828	A2	01 October 2008
				GB	0918569	D0	09 December 2009
				GB	2461453	A	06 January 2010
				GB	0918582	D0	09 December 2009
				GB	2460804	A	16 December 2009
				GB	0918583	D0	09 December 2009
				GB	2460991	A	23 December 2009
				GB	2460991	A8	20 January 2010
				GB	201008893	D0	14 July 2010
				GB	2468070	A	25 August 2010
				GB	2468070	B	06 June 2012
				KR	20090044978	A	07 May 2009
				KR	101598531	B1	29 February 2016
				US	2008242947	A1	02 October 2008
				US	2008242948	A1	02 October 2008
				US	2008242949	A1	02 October 2008
				US	2008242950	A1	02 October 2008
				US	2008242951	A1	02 October 2008
				US	2008242952	A1	02 October 2008
				US	2008243005	A1	02 October 2008
				US	2008243543	A1	02 October 2008
				US	2008287821	A1	20 November 2008
				US	2008319276	A1	25 December 2008
				US	2009005653	A1	01 January 2009
				US	2009005654	A1	01 January 2009
				US	2009018407	A1	15 January 2009
				US	2009024050	A1	22 January 2009
				US	2009112616	A1	30 April 2009
				US	2009112617	A1	30 April 2009
				US	8065240	B2	22 November 2011
				US	2009112620	A1	30 April 2009
				US	2009112621	A1	30 April 2009
				US	2009118593	A1	07 May 2009
				US	2009119154	A1	07 May 2009
				US	2009132275	A1	21 May 2009
				US	2012164613	A1	28 June 2012
				US	2018254103	A1	06 September 2018
				US	2019239789	A1	08 August 2019
				US	2020329969	A1	22 October 2020
				WO	2008121371	A2	09 October 2008
				WO	2008121371	A3	27 November 2008
				WO	2008121399	A2	09 October 2008
				WO	2008121399	A3	20 November 2008
				WO	2008121401	A2	09 October 2008
				WO	2008121401	A3	27 November 2008
				WO	2008143908	A2	27 November 2008
				WO	2008143908	A3	22 January 2009
				WO	2008143941	A1	27 November 2008
				WO	2009058164	A1	07 May 2009

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/IL2024/050611

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CA	3193776	A1	31 March 2022	CA	3193776	A1	31 March 2022
				AU	2021347379	A1	25 May 2023
				AU	2021347379	A9	08 August 2024
				CN	116829050	A	29 September 2023
				EP	4216807	A1	02 August 2023
				JP	2023544550	A	24 October 2023
				KR	20230137869	A	05 October 2023
				US	2023255564	A1	17 August 2023
				WO	2022067189	A1	31 March 2022