United States Patent Application Publication 20250262461
Kind Code A1
Publication Date August 21, 2025
Inventor(s) DAFT; Christopher

ULTRASOUND NEUROMODULATION GUIDED BY ARTIFICIAL INTELLIGENCE

Abstract

A medical ultrasonic system employs an artificial intelligence large vision model (LVM) trained on a database of target anatomy. The system employs an ultrasound probe and a neuromodulation beamformer plus additional instrumentation, all interfaced directly or indirectly to the domain specific large vision model (DSLVM). The system is configured to operate using volume ultrasound imaging or slice ultrasound imaging of the target anatomy. Examples are described for the cranial anatomy.

Inventors: **DAFT; Christopher (Tucson, AZ)**

Applicant: **Sanmai Technologies, PBC** (Sunnyvale, CA)

Family ID: **1000008319617**

Appl. No.: **18/959039**

Filed: **November 25, 2024**

Related U.S. Application Data

us-provisional-application US 63554004 20240215

Publication Classification

Int. Cl.: **A61N7/00** (20060101); **G06V10/774** (20220101); **G16H20/40** (20180101); **G16H30/40** (20180101)

U.S. Cl.:

CPC **A61N7/00** (20130101); **G06V10/774** (20220101); **G16H20/40** (20180101); **G16H30/40** (20180101); A61N2007/0026 (20130101); A61N2007/0078 (20130101); G06V2201/03

(20220101)

Background/Summary

CROSS-REFERENCE TO RELATED APPLICATIONS [0001] The present application claims the benefit of U.S. Provisional Application titled "ULTRASOUND STIMULATION GUIDED BY ARTIFICIAL INTELLIGENCE", filed on Feb. 15, 2024, and having Ser. No. 63/554,004. The subject matter of this related application is hereby incorporated herein by reference.

TECHNICAL FIELD

[0002] This invention relates to a medical ultrasound system that employs a domain specific large vision model (DSLVM) trained on a database of target anatomy, and in particular to a system adaptable to using either volume or slice imaging.

BACKGROUND

[0003] In a medical ultrasound system, the critical functions of targeting and neuromodulation control have previously been performed by human operators aided by computer vision algorithms. In the present invention, new methods for performing these functions are proposed.

Description

BRIEF DESCRIPTION OF THE DRAWINGS

[0004] FIG. **1** illustrates training of a device specific large vision model (DSLVM) using data from an MRI database for the target anatomy.

[0005] FIG. **2** illustrates a DSLVM operating in inference mode.

[0006] FIG. **3** illustrates an ultrasound system centered on a large vision model, performing inference using a volume-scanning ultrasound probe.

[0007] FIG. **4** illustrates a slice geometry comprising an azimuth probe element having elevation angle greater than azimuth angle.

[0008] FIG. **5** illustrates an alternative ultrasound system centered on a large vision model, performing inference using a slice probe.

DETAILED DESCRIPTION OF THE INVENTION

[0009] Two medical ultrasound systems are described, each system centered on an artificial intelligence (AI) system which performs the critical functions of targeting and neuromodulation control. Large Vision Models (LVMs), a form of AI, have evolved from Large Language Models (LLMs) and are oriented towards images rather than text. Examples of early LVMs include VGGNet, GoogleNet, and ResNet. In this application, Domain Specific Large Vision Models (DSLVMs) are directed at medical ultrasound systems.

[0010] The two ultrasound systems differ in the types of probe and data acquisition systems used. The first to be described (see FIG. **3**) uses a 2D or matrix probe which can acquire volume data. In it, the AI's encoded knowledge of a large database of MRI scans relates the real-time volume data being acquired by the probe to an estimated MRI. The system of FIG. **3** does not require a pre-procedure MRI scan.

[0011] The second system (see FIG. **5**) uses a 1D or other probe which produces slice data by electronically scanning in one dimension. It uses the AI to relate the real-time slice data being acquired by the probe to a pre-procedure structural MRI scan.

[0012] Three architectures have emerged for DSLVMs: attention-based, convolutional, and multi-layer perceptron. Firstly, in the prior art, the transformer architecture became a building block for constructing DSLVMs. Examples include Vision Transformer (ViT, Google Brain), Swin Transformer (Microsoft), and VideoMAE (Tencent). These DSLVMs undergo pre-training using extensive image datasets, enabling them to capture image content and extract semantic information. Meta's Segment Anything Model (SAM) operates as an image segmentation model which can be prompted.

[0013] Secondly, convolutional neural networks (CNNs) historically stood out due to their computational efficiency over multilayer perceptrons (MLPs). CNNs outperform traditional algorithms for almost all computer vision tasks, such as object detection, segmentation, de-mosaicing, super-resolution, and deblurring. The CNN architecture typically comprises alternate convolutional and pooling layers with several fully connected layers behind.

[0014] Thirdly, Google Brain developed the MLP-mixer architecture, another new computer vision paradigm. It uses neither the attention mechanism of transformers, nor convolution operations.

[0015] Disclosed herein is a DSLVM-centric ultrasound treatment system where guidance of the neuromodulation is based on real-time ultrasound plus anatomical knowledge encapsulated in a DSLVM. The target anatomy may be a structure with a subject's brain, or any other anatomy. The DSLVM may be built from any of the architectures described—a transformer, a CNN or an MLP-mixer, or another approach.

[0016] FIG. **1** illustrates a training system **10** for a DSLVM **14**, employing volume anatomical database **11** to illustrate the range of anatomy in the clinical area of interest. Predicted target locations **17** corresponding to a trained DSLVM may be obtained using transfer learning from a similar model trained using MRI and computerized tomography (CT), such as V-Net, (reference 4). Alternatively, it may be trained from scratch using MRI, CT or other volumetric scans. In either case, many ultrasound training datasets are created from each MRI scan, each having a different transducer type, such as a slice probe with a 1D line of elements or a volume probe with a 2D matrix of elements. Different probe geometries and center frequencies generate further ultrasound synthetic training data from the MRI scan. Simulations for model training include many probe locations on the scalp.

[0017] The DSLVM is trained to comprehend the relationship between ultrasound slice or volume data and a volumetric map of the target anatomy. Synthetic ultrasound data indicated in the box “Ultrasound volume or slice data” **13***d* can be reliably obtained from MRI or other volume scanning modalities in a few steps. For example, this may be done by first creating a synthetic CT image using established deep learning methods (reference 7) and then estimating the acoustic properties **13***a* of the target anatomy from the synthetic CT (reference 8). Alternatively, density and speed of sound maps **13***a* may be computed from a database of CT scans. Other routes from volume medical imaging modalities to acoustic property data sets are also possible. With these data, a model such as a pseudospectral simulation **13***c* can reliably model ultrasonic propagation and scattering and create simulated ultrasonic returns to probe **13***b* to train the model. Transducer type and position obtained from the Ultrasound Probe Specification **13***b* are also input to the pseudospectral simulation **13***c*. Other types of ultrasonic simulation algorithms such as finite-element or finite-difference models may also be used. Labeling or visual prompting **15** may be applied to volume anatomical database **11** but this may not be required since the DSLVM may be capable of identifying the relevant structures without labels. The output of ultrasound simulation **13***c*, either in slice or volume form **13***d*, is fed to the DSLVM.

[0018] FIG. **2** illustrates a system **20** comprising a DSLVM **21** operating in

[0019] inference mode, as would be the case in a product. Ultrasound probe **22** produces real-time ultrasound data **23** in various modes which may include: tissue harmonic imaging (THI), Power Doppler (PD, blood vessels), Strain (tissue stiffness), or a 2-transducer mode that transmits using one aperture and receives using a second aperture. Real-time ultrasound data **23** may be augmented with optical or LIDAR data about the subject's head geometry from a camera swept around the target anatomy (not shown). The operator or clinician specifies the name of the structure to be treated **24** and the DSLVM outputs target coordinates **25** to Real-time ultrasound controller **26** which, after computation of beam formation coefficients, drives the Neuromodulation Beamformer

27. Calculating the beam formation coefficients of an ultrasonic field which matches coordinates bounding the target structure(s) is a well-studied area in therapeutic ultrasound (reference 9).

[0020] DSLVM **21** also creates confirmatory outputs **28**. These show the operator aspects of the treatment such as a real-time anatomical illustration with the treatment volume(s) highlighted. The DSLVM is trained to produce an MRI-style anatomical estimate in the ultrasound probe's coordinate system from ultrasound data acquired at any position or angle.

[0021] Using the background described in reference to FIGS. **1** and **2**, further details are provided in each of two embodiments of the present disclosure: (i) the volume ultrasound approach and (ii) the slice ultrasound approach. The volume ultrasound approach to be described in reference to FIG. **3** is superior; it employs a probe comprising a 2D matrix of transducer elements and an application-specific integrated circuit (ASIC) to process the high information rate from the matrix array (reference 10). The slice ultrasound approach to be described in reference to FIGS. **4** and **5** produces far less real-time anatomical data which is a limitation. But the hardware system can use traditional, low-cost transducer manufacturing, which can be an advantage in practice when making a product. It has lower hardware technical risk but relies more on the capability of the DSLVM than does the implementation of FIG. **3**.

[0022] FIG. **3** illustrates the volume ultrasound embodiment **30** of the present disclosure. It shows how a volume-scanning matrix probe **31***b* and EEG or MEG sensors **31***a* provide real-time inputs **31** to the DSLVM. The ultrasound volume data prompts the model **32**, whose outputs **33** can be used to program a delay and amplitude controlling beamformer to provide ultrasonic neuromodulation to a target anatomy. An AI processor **32***a* within the domain specific large vision model **32** produces ultrasound reachable target outline coordinate points **33** used to create beamformer coefficients **34**, and slow-time simulation control **35** derived from EEG/MEG **31***a* that provides timing information to the neuromodulation hardware **36**. In an embodiment AI processor **32***a* comprises a network of GPU and CPU processors. DSLVM **32** acts like the kernel of an operating system; its function is analogous to the way the Linux kernel interacts with peripherals attached to its CPU. Here, devices analogous to peripherals in a Linux system include: [0023] An electroencephalographic (EEG) **31***a* or magnetoencephalographic (MEG) array of sensors which can provide information about the response to neuromodulation of the target anatomy. [0024] A matrix array of ultrasound transducers and beamforming electronics capable of volume acquisition **31***b* of anatomical and blood flow data. [0025] Desired targeting data **38** specified by a clinician by either: [0026] A text string such as “posterior cingulate cortex and amygdala,” or: [0027] A set of points marked by the clinician on a pre-procedure MRI scan, if available. These points define a treatment volume in the coordinate system of the MRI, which differs from the coordinate system of the neuromodulation transducer. [0028] An operator display **37** provides access to several data sources informing a clinician of the progress made by the system towards reliable targeting of the neuromodulation. This display may include: [0029] Volume ultrasound data **37***a* being received from the imaging/guidance hardware. [0030] A “fuel gauge” **37***b* indicating to the clinician the DSLVM's confidence about the usefulness of the current probe location and orientation for the neuromodulation task. [0031] Confirmatory outputs **37***c* providing further data to the clinician on the overall credibility of the DSLVM's interpretation of the target anatomy revealed by the real-time ultrasound. [0032] Neuromodulation hardware **36**, comprising a probe, drive electronics creating appropriate neuromodulation drive voltages to be applied to each probe element, and a beamformer which translates a description of the volume to be treated to parameters controlling the amplitude and time delays of the voltages applied to the probe elements which deliver the treatment specified by the DSLVM. [0033] Slow-time neuromodulation control **35**, informed by the EEG or MEG information supplied to the LVM. These data may define when neuromodulation is applied relative to the activity in the target anatomy.

[0034] The behavior of the neuromodulation hardware **36** is controlled in two ways. Steering, focusing and other control of the spatial extent of the beam are accomplished by selection of beamforming coefficients **34**. Temporal control of the neuromodulation to synchronize with aspects of physiological activity is achieved by the separate slow-time neuromodulation control **35** that is output from the DSLVM **32**.

[0035] In an embodiment of the present disclosure a use-case for the volume ultrasound approach of FIG. **3** includes the following steps, wherein the target anatomy is the cranial region of a subject: [0036] Specify the treatment region: this can be either: [0037] Part of prescription from the prescribing doctor. [0038] Wellness applications which may be accessible without prescription [0039] Place the probe on the subject's skull. [0040] The DSLVM converts the volume ultrasound data into anatomy and labels the treatment area. [0041] If the target anatomy probe proves to be in a poor location for treatment, the operator display (in a clinical environment) or the subject's phone (during at-home use) may show how to reposition it. [0042] When the probe location is adequate for treatment, treatment starts.

[0043] In an embodiment of the present disclosure FIG. **4** illustrates a 1.75D probe **40** comprising an array **41** of probe elements having elements comprising an elevation dimension **42** larger than the azimuth dimension **43**. In an exemplary application, array **40** is used to steer and focus ultrasonic waves in a transcranial ultrasound (TUS) procedure. Four azimuth probe elements **44** are shown as one column of array **41**. The 1.75D arrangement economizes on the total number of probe elements by only contemplating steering in the azimuth dimension. Efficacy of the accompanying tFUS procedure is achieved with this reduced number of elements using the system of FIG. **5**. The smaller element count allows a standard cable to connect the probe array **41** to a signal processor. The probe can be built with traditional, low-cost manufacturing techniques employing bulk PZT. The size of the elements is insufficient to support steering in elevation, but the element dimensions are sufficient to perform aberration correction in elevation as well as in azimuth. The data created using probe array **41** may be called “slice data” which has been corrected for skull aberration.

[0044] In an embodiment of the present disclosure FIG. **5** illustrates the slice ultrasound approach **50** wherein the target anatomy is a subject's cranial region. A 1.75D probe **40** is used as a real-time input **51**, connected to slice ultrasound hardware **51***a*, and then to DSLVM **52**. The DSLVM may also receive EEG data **51***b* (or magnetoencephalogram data, MEG data), and probe orientation data (produced, for example, by MEMs gyro **51***c*. Fixed inputs **53** include pre-procedure MRI **53***a* (which is volume data) and a targeting command (specified as text or as a set of points marked by the operator on the MRI) **53***b*. An AI processor **52***a* within the DSLVM **52** outputs to operator display **54** the next probe position/orientation **54***a*, data for a “fuel gauge” showing how complete the acquisition **54***b* is, and a second fuel gauge showing a confidence level in the treatment coordinates **54***c*, and a real-time head model with treatment outline **54***d*. In an embodiment AI processor **52***a* comprises a network of GPU and CPU processors. The head model can be thought of as a rotation and translation of the pre-procedure MRI data. The real-time head model **54***d* is overlaid with treatment outline data from coordinate points **56**. Beamformer coefficients define the steering and focusing of neuromodulation hardware **59**, and are produced from coordinate points **56**. Slow-time simulation control **58**, for example to coordinate the timing of neuromodulation beams with physiological events sensed by EEG or MEG **51***b*, may also control neuromodulation hardware **59**. Coordinate points may include fiducial structures or blood vessel locations within the target outline.

[0045] In an embodiment of the present disclosure a use-case for the slice ultrasound approach depicted in FIG. **5**, wherein the target anatomy is a subject's cranial region, includes the following steps: [0046] The DSLVM reads the pre-procedure MRI volume. [0047] The operator defines the structure to be treated either as text or by annotating the MRI volume. [0048] The operator starts to scan the subject's head with the ultrasound probe [0049] The operator moves the probe around on the subject's head, acquiring slices for DSLVM to use. This is necessary because probe **40** is not acquiring volume data, in contrast to the situation in the system of FIG. **3**. [0050] The operator views the MRI **54***d* translated into the probe's coordinate system on display **54**, wherein the

DSLVM produces an MRI-to-live-probe coordinate transformation. [0051] The operator monitors the completeness of data acquired so far by observing fuel gauge **54***b*. as they move the probe on the subject's head. [0052] The DSLVM **52** provides directions **54***a* informing the operator of the best way to move the probe to complete the data acquisition. [0053] Once fuel gauge **54***b* indicates that enough data has been acquired, probe position indicator **54***a* shows the ideal position for treatment.

[0054] The operator fixes the probe at the indicated position and reviews the fuel gauge **54***c* to decide whether the acquired data and probe position provide confidence in the treatment setup.

[0055] If the display data is approved by the clinical operator, they start the treatment

[0056] Several differences are apparent between FIGS. **3** and **5**, reflecting the increased complexity in achieving full guidance control with a guidance array capable only of producing electronically scanned slice data. Specifically: [0057] An additional real-time input may be provided to the DSLVM by a MEMS gyro **51***c* which reports the orientation of the probe. Similarly, a MEMS accelerometer (not shown) may also be mounted on the probe to report its motion. [0058] Key to the operation of the apparatus of FIG. **5** is the pre-procedure MRI **53***a* of the subject's head. This provides per-subject anatomical information to correlate with the real-time ultrasound slice data. [0059] The operator display includes a real-time head model with a treatment outline **54***d*. The DSLVM creates the treatment volume outline from a textual prompt or from a set of points marked on the MRI by the clinician. The DSLVM also determines how to rotate and translate the MRI to appear in the coordinate system of the neuromodulation transducer. This operation is simplified if the neuromodulation probe is the same as, or is rigidly attached to, the guidance probe. [0060] As will be clear from the previously described method, the slice transducer may need to image the head from multiple viewpoints before the DSLVM has enough confidence that the probe position, and orientation can be moved to a suitable plane for neuromodulation of the target. [0061] Another part of the display **54***a* shows the head outline and the location of the probe on it, together with arrows recommending translation and rotation for a new acquisition which efficiently provides an adequate concept of the brain's volume anatomy for targeting to succeed. [0062] Another display element **54***b* shows the DSLVM's view of the completeness of acquisition. With a slice acquisition, several views will need to be obtained before the DSLVM can have confidence in mapping the pre-procedure MRI to the live ultrasound. [0063] Once enough slices have been acquired for reliable targeting, the DSLVM specifies to the operator the optimal position of the probe for neuromodulation treatment. Again, arrows are helpful in showing the operator how to move from an initial location to the best location and orientation.

[0064] Various embodiments of the present disclosure are described in the following clauses. Although the following clauses describe some embodiments of the present disclosure, other embodiments of the present disclosure are also set forth above.

[0065] 1. In some embodiments, a system for training a large vision model comprises a specification for an ultrasound probe, a database of volumetric anatomical scans, a prompt string defining an anatomical target, at least one computing device, and an application executed by the at least one computing device that, when executed, causes the at least one computing device to at least train a large vision model where the inputs comprise the prompt string and data obtained by simulating the field produced by the specified ultrasonic probe and one of the volumetric anatomical scans, and the trained output produces weights that generate a set of target locations.

[0066] 2. The system of clause 1 wherein the specified ultrasonic probe comprises a two-dimensional array of elements.

[0067] 3. The system of clauses 1 or 2 wherein the specified ultrasonic probe comprises a one-dimensional array of elements.

[0068] 4. The system of any of clauses 1-3 wherein the target anatomy is a structure within a human brain.

[0069] 5. The system of any of clauses 1-3 wherein the large vision model is trained on a database

of cranial data.
[0070] 6. In some embodiments, a system comprises an ultrasonic probe, a neuromodulation beamformer coupled to the ultrasonic probe, at least one computing device, and an application executed by the at least one computing device that, when executed, causes the at least one computing device to at least effect inference using an AI processor within a large vision model (LVM) based on a database of target anatomy and data obtained by at least one of the ultrasonic probe or the neuromodulation beamformer, and produce coordinates of the target anatomy.
[0071] 7. The system of clause 6 wherein the ultrasonic probe comprises a two-dimensional array of elements.
[0072] 8. The system of claim **6** wherein the ultrasonic probe comprises a one-dimensional array of elements.
[0073] 9. The system of claim **6** further comprising a prompt to indicate the target structure.
[0074] 10. The system of claim **6** further comprising a volume scan of the head acquired prior to the procedure.
[0075] 11. The system of claim **6** further comprising beam formation electronics for neuromodulation.
[0076] 12. The system of claim **6** wherein each generator comprises a rectangular element having its elevation larger than its azimuth.
[0077] 13. The system of claim **6** wherein the large vision model identifies a best match out of a database of MRI volumes.
[0078] 14. The system of claim **6** wherein items are displayed to a clinician confirming correct targeting.
[0079] 15. The system of claim **6** wherein the large vision model is guided by ultrasonically measured coordinates of at least one of fiducial structures or blood vessel locations within the target outline.
[0080] 16. The system of any of clauses 6-15, and a second fuel gauge that shows a targeting confidence level.
[0081] 17. The system of any of clauses 6-16, the user interface comprising at least one of a Grayscale B mode, a Tissue Harmonic Imaging (THI) mode, a color flow mapping mode, a power doppler mode, or a strain imaging mode.
[0082] 18. The system of claim **6** wherein the user interface comprises additional data provided using a body scan produced by a camera or LIDAR device.
[0083] 19. The system of claim **6** wherein the target anatomy is a structure within a human brain.
[0084] 20. The system of claim **6** wherein the timing of neuromodulation emissions is controlled by a slow-time modulation control provided by the LVM.
[0085] While the preceding is directed to embodiments of the present disclosure, other and further embodiments of the disclosure may be devised without departing from the basic scope thereof, and the scope thereof is determined by the claims that follow.

Claims

1. A system for training a large vision model comprising: a specification for an ultrasound probe; a database of volumetric anatomical scans; a prompt string defining an anatomical target; at least one computing device; and an application executed by the at least one computing device that, when executed, causes the at least one computing device to at least: train a large vision model where the inputs comprise the prompt string and data obtained by simulating the field produced by the specified ultrasonic probe and one of the volumetric anatomical scans; and the trained output produces weights that generate a set of target locations.
2. The system of claim 1 wherein the specified ultrasonic probe comprises a two-dimensional array of elements.

3. The system of claim 1 wherein the specified ultrasonic probe comprises a one-dimensional array of elements.
4. The system of claim 1 wherein the target anatomy is a structure within a human brain.
5. The system of claim 1 wherein the large vision model is trained on a database of cranial data.
6. A system comprising: an ultrasonic probe; a neuromodulation beamformer coupled to the ultrasonic probe; at least one computing device; and an application executed by the at least one computing device that, when executed, causes the at least one computing device to at least: effect inference using an AI processor within a large vision model (LVM) based on a database of target anatomy and data obtained by at least one of the ultrasonic probe or the neuromodulation beamformer; and produce coordinates of the target anatomy.
7. The system of claim 6 wherein the ultrasonic probe comprises a two-dimensional array of elements.
8. The system of claim 6 wherein the ultrasonic probe comprises a one-dimensional array of elements.
9. The system of claim 6 further comprising a prompt to indicate the target structure.
10. The system of claim 6 further comprising a volume scan of the head acquired prior to the procedure.
11. The system of claim 6 further comprising beam formation electronics for neuromodulation.
12. The system of claim 6 wherein each generator comprises a rectangular element having its elevation larger than its azimuth.
13. The system of claim 6 wherein the large vision model identifies a best match out of a database of MRI volumes.
14. The system of claim 6 wherein items are displayed to a clinician confirming correct targeting.
15. The system of claim 6 wherein the large vision model is guided by ultrasonically measured coordinates of at least one of fiducial structures or blood vessel locations within the target outline.
16. The system of claim 6 wherein the user interface further comprises a first fuel gauge that shows the completeness of acquisition of anatomical information, and a second fuel gauge that shows a targeting confidence level.
17. The system of claim 6 wherein the application generates a user interface, the user interface comprising at least one of a Grayscale B mode, a Tissue Harmonic Imaging (THI) mode, a color flow mapping mode, a power doppler mode, or a strain imaging mode.
18. The system of claim 6 wherein the user interface comprises additional data provided using a body scan produced by a camera or LIDAR device.
19. The system of claim 6 wherein the target anatomy is a structure within a human brain.
20. The system of claim 6 wherein the timing of neuromodulation emissions is controlled by a slow-time modulation control provided by the LVM.