



OPEN ACCESS

EDITED BY

Walid Kamal,
University of Sharjah, United Arab Emirates

REVIEWED BY

Kalyana Chakravarthy Bairapareddy,
University of Sharjah, United Arab Emirates
Ahmed Abdelmoniem Ibrahim,
University of Hail, Saudi Arabia
Fatma Refaat Ahmed,
University of Sharjah, United Arab Emirates

*CORRESPONDENCE

Pedro Guimarães
✉ pedro.guimaraes@icnas.uc.pt

RECEIVED 13 October 2023

ACCEPTED 14 December 2023

PUBLISHED 05 January 2024

CITATION

Guimarães P, Serranho P, Duarte JV,
Crisóstomo J, Moreno C, Gomes L,
Bernardes R and Castelo-Branco M (2024)
The hemodynamic response function as a
type 2 diabetes biomarker: a data-driven
approach.
Front. Neuroinform. 17:1321178.
doi: 10.3389/fninf.2023.1321178

COPYRIGHT

© 2024 Guimarães, Serranho, Duarte,
Crisóstomo, Moreno, Gomes, Bernardes and
Castelo-Branco. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication
in this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

The hemodynamic response function as a type 2 diabetes biomarker: a data-driven approach

Pedro Guimarães^{1*}, Pedro Serranho^{1,2}, João V. Duarte^{1,3},
Joana Crisóstomo¹, Carolina Moreno⁴, Leonor Gomes⁴,
Rui Bernardes^{1,5} and Miguel Castelo-Branco^{1,5}

¹University of Coimbra, Coimbra Institute for Biomedical Imaging and Translational Research (CIBIT), Institute for Nuclear Sciences Applied to Health (ICNAS), Coimbra, Portugal,

²Department of Sciences and Technology, Universidade Aberta, Lisbon, Portugal, ³University of Coimbra, Faculty of Medicine (FMUC), Coimbra, Portugal, ⁴Department of Endocrinology, University Hospital of Coimbra (CHUC), Coimbra, Portugal, ⁵University of Coimbra, Clinical Academic Center of Coimbra (CACC), Faculty of Medicine (FMUC), Coimbra, Portugal

Introduction: There is a need to better understand the neurophysiological changes associated with early brain dysfunction in Type 2 diabetes mellitus (T2DM) before vascular or structural lesions. Our aim was to use a novel unbiased data-driven approach to detect and characterize hemodynamic response function (HRF) alterations in T2DM patients, focusing on their potential as biomarkers.

Methods: We meshed task-based event-related (visual speed discrimination) functional magnetic resonance imaging with DL to show, from an unbiased perspective, that T2DM patients' blood-oxygen-level dependent response is altered. Relevance analysis determined which brain regions were more important for discrimination. We combined explainability with deconvolution generalized linear model to provide a more accurate picture of the nature of the neural changes.

Results: The proposed approach to discriminate T2DM patients achieved up to 95% accuracy. Higher performance was achieved at higher stimulus (speed) contrast, showing a direct relationship with stimulus properties, and in the hemispherically dominant left visual hemifield, demonstrating biological interpretability. Differences are explained by physiological asymmetries in cortical spatial processing (right hemisphere dominance) and larger neural signal-to-noise ratios related to stimulus contrast. Relevance analysis revealed the most important regions for discrimination, such as extrastriate visual cortex, parietal cortex, and insula. These are disease/task related, providing additional evidence for pathophysiological significance. Our data-driven design allowed us to compute the unbiased HRF without assumptions.

Conclusion: We can accurately differentiate T2DM patients using a data-driven classification of the HRF. HRF differences hold promise as biomarkers and could contribute to a deeper understanding of neurophysiological changes associated with T2DM.

KEYWORDS

type 2 diabetes, functional magnetic resonance imaging, neuroimaging, deep learning, hemodynamic response

1 Introduction

Type 2 diabetes mellitus (T2DM) is a prevalent chronic metabolic disorder with an enormous global impact. In 2021, approximately 537 million adults aged 20–79 had diabetes, a number expected to reach 783 million by 2045 due to rising obesity and sedentary lifestyles (IDF Diabetes Atlas, 2021). It is a significant burden on healthcare systems, as disease management and complications can be costly.

T2DM is a systemic disease impacting multiple organs and tissues. In the brain, it is associated with an increased risk of function loss and long-term cognitive impairment (Biessels et al., 2014; Biessels and Despa, 2018). Patients are at increased risk for cerebrovascular disease, dementia, and depression, among other neuropathies (Biessels et al., 2014; Bancks et al., 2017; Biessels and Despa, 2018). To support research and preventive therapies development, it's important to study the neurophysiological changes associated with early brain dysfunction in T2DM subjects before overt vascular or structural lesions.

Although there is no consensus on their full extent and mechanisms, T2DM-related changes in the brain have been documented (Biessels and Reijmer, 2014; Chen et al., 2015; Duarte et al., 2015; Crisóstomo et al., 2021; Zhang et al., 2022; Duarte et al., 2023), with most literature focused on structural differences. Among the observed changes, magnetic resonance imaging (MRI) shows both local and general abnormalities in T2DM subjects, with a wide prevalence of cortical and subcortical gray matter atrophy, as shown by volumetric analysis at the whole-brain level and voxel-based morphometry (Moran et al., 2013; Moheet et al., 2015). Cortical thickness was found to be reduced in T2DM patients compared to controls (Chen et al., 2015; Crisóstomo et al., 2021). Differences in cortical gyration were also associated with T2DM, mainly in cortical sensory areas (Crisóstomo et al., 2021).

Functional brain changes have also been described. However, these reports are rather inconsistent. Both task-based and resting state studies of blood flow and oxygenation level have been used to uncover neuropathological mechanisms in T2DM. However, resting state studies are much more common. Cerebral blood flow was found to be altered in T2DM patients in Ryan et al. (2014), but with a larger sample size, no statistically significant differences were found when accounting for brain volume (Tiehuis et al., 2008).

Functional magnetic resonance imaging (fMRI) with blood oxygen level-dependent (BOLD) contrast can indirectly quantify neural activity. Widespread abnormal activity has been reported in T2DM (Chhatwal and Sperling, 2012; Xia et al., 2013; Cui et al., 2014; Liu et al., 2018, 2019; Xiong et al., 2020). However, the pathophysiological mechanisms leading to structural changes, behavioral impairment, and brain dysfunction are still not well understood. Deconvolution analysis was used to examine the fMRI response profile in Duarte et al. (2015, 2023). Differences were found between controls and lesion-free T2DM patients, in this approach which considered the hemodynamic response function of capillary

networks. However, data driven approaches are needed to validate these findings.

Brain imaging has become ubiquitous in neuroscience and mostly with model-driven approaches. In parallel, deep learning (DL) has shown incredible potential in medical image analysis (Litjens et al., 2017). Unsurprisingly, the neuroimaging and DL fusion has been promptly adopted by research groups around the world (Litjens et al., 2017; Zaharchuk et al., 2018; Warren and Moustafa, 2023). fMRI-based research has seen a surge in neural network (NN) applications. In Sarraf et al. (2016), scans from controls and Alzheimer's disease patients were concatenated into 2D images and classified using convolutional NNs (CNNs). In Horikawa and Kamitani (2017), CNNs were used to decode seen and imagined objects. DL has also been used with resting state fMRI to classify chronic pain conditions (Santana et al., 2019), mild cognitive impairment (Suk et al., 2016), or schizophrenia (Kim et al., 2016).

In this work, we use a combination of task-based fMRI and DL to investigate whether T2DM patients have an altered BOLD response compared to age- and gender-matched controls. In contrast to typical fMRI studies, we used an event-related design, as this is the only design suitable to study an altered hemodynamic response function (HRF). This is of particular importance in the case of diabetes, where the assumption of a standard HRF is not guaranteed. We explore the explainability of DL methods to improve decision transparency and combine novel DL-based methods and deconvolution generalized linear model (GLM) analysis to reveal the true HRF in T2DM subjects without any assumptions.

2 Materials and methods

2.1 Ethics statement

This study was approved by the Ethics Committee of the Faculty of Medicine of the University of Coimbra (ethical approval number CE-059-2019). All participants provided informed written consent. The Helsinki Declaration of 1975 (revised in 1983) guidelines were followed throughout.

2.2 Participants

Recruited T2DM patients were diagnosed at the Endocrinology Department of the University of Coimbra Hospital following the World Health Organization criteria (HbA1c and fasting plasma glucose levels) and were drawn from the department's clinical population. Inclusion criteria were: (1) between 40–75 years of age, (2) T2DM for at least 1 year prior, and (3) informed consent. Diagnosis and severity of diabetic retinopathy were determined by hospital experts based on a technical report according to the Early Treatment Diabetic Retinopathy Study guidelines.

Control subjects were recruited from the general population of the Hospital or University staff. Inclusion criteria were: (1) between 40–75 years of age, (2) T2DM diagnosis was ruled out (based on HbA1c and fasting plasma glucose levels), and (3) informed consent.

Eye and hand dominance were assessed using the hole-in-the-card test (Dolman method) and the Edinburgh inventory, respectively. All participants had either normal or corrected-to-normal vision and no history of any neurological or psychiatric disease. Fundus photography and optical coherence tomography were used to confirm the visual system's health.

A neuroradiologist assessed standard clinical 2D-FLAIR and 3D-SPACE T2-weighted scans for the presence of white matter hyperintensities. Participants with vascular or structural lesions or any confounding changes were excluded. Participants with a maximum head movement of more than 3 mm during acquisition, as detected by motion-correction, were also excluded.

After exclusion, 51 T2DM and 67 control subjects were selected. Participants were randomly divided into two datasets: Dataset 1 (DS1), containing 80% of all participants, and Dataset 2 (DS2), containing the remaining 20%. Two constraints were imposed on the split: (1) the same proportion of T2DM subjects was forced in both sets; (2) for DS2, the age and gender distributions between T2DM and control subjects were forcibly matched, i.e., we randomly selected either a T2DM or control subject and then found the closest match in terms of gender and age from the opposite class.

2.3 Experimental design and imaging

All participants were imaged at the facilities of the Portuguese Brain Imaging Network in a 3 Tesla scanner (Magnetom TIM Trio, phased array 12-channel birdcage head coil; Siemens, Munich, Germany). Imaging comprised magnetic resonance (MR) anatomical and functional scans. Three-dimensional anatomical MPRAGE (magnetization-prepared rapid gradient echo) scans were acquired using a standard T1-weighted gradient echo (GE) pulse sequence (2,530 ms repetition time (TR); 3.42 ms echo time (TE); 1,100 ms inversion time (TI); 7° flip angle; 176 slices; 1 × 1 × 1 mm voxel size; 256 mm field of view (FOV)).

The functional series consisted of a run of 116 GE, EPI scans (2,500 ms TR; 30 ms TE; 90° flip angle; 36 interleaved slices, 3 × 3 × 3 mm voxel size; 256 mm FOV) in an event-related design stimulation paradigm with measurement of the BOLD signal.

2.3.1 Event-related task: stimuli description

Subjects were asked to compare the speed of two white dots, a reference, and a target dot, with 22.9 cd/m^2 mean luminance moving on a gray background (9.39 cd/m^2 mean luminance) and to select the fastest of the two.

Stimuli setup was prepared in MATLAB using the Psychophysics Toolbox (David Brainard and Denis Pelli, <http://www.psychtoolbox.org>) and projected with an LCD projector (Avotec Real Eye Silent Vision 6,011, 1,024 × 768 pixels, 60 Hz refresh-rate; Avotec Incorporated, Stuart, FL, USA) onto a screen pad positioned in the bore at 163 cm (22.62° × 17.06° image size; mirror 50 cm from the screen). Stimuli were viewed with one eye while the other eye was covered with an opaque patch.

Two dots (0.22° × 0.22° dot size, 0.33° × 0.33° fixation-cross size) were presented, one in each visual hemifield, randomly distributed at

7.5° eccentricity along the horizontal meridian, moving back and forth along 2° path in a pseudo-random linear trajectory (between 0 and 180°). The subject selects the faster moving dot by pressing a button on a Cedrus Lumina LP-400, LU400 PAIR response box (Cedrus Corporation, San Pedro, CA, USA).

2.3.2 Speed threshold discrimination task

Prior to functional acquisition, each subject's speed discrimination threshold was determined, i.e., the limit delta between a reference and target dot at which each participant was able to correctly discriminate the faster dot. This was done inside the scanner with the dominant eye. The dots were in view for 400 ms. Successive trials were performed with decreasing delta. The speed of the target dot on successive trials was determined using a descending logarithmic staircase. The staircase had six reversals (two practice/four experimental). It allowed us to compute speeds from 24°/s (initial target speed) to 5°/s (reference speed) using a step value ranging from 1–0.05 dB. The threshold was estimated using the average of the last four reversals.

2.3.3 Stimulation conditions

Functional acquisition for the event-related design stimulation paradigm included two conditions, designated threshold and sub-maximum, each presented 10 times per hemifield to increase statistical power. For the threshold condition the target dot moved at the reference speed plus the discrimination threshold (individually determined); and for the sub-maximum condition the target dot moved at the reference speed plus three times the same threshold. Note that in both conditions the reference dot moved at the reference speed (5°/s).

Each stimulation period lasted 400 ms, and the baseline fixation period was randomly assigned to be either 4,600, 7,100, or 9,600 ms. Participants were instructed to maintain fixation on the fixation-cross. The stimulation and imaging protocol have been further described in Duarte et al. (2015).

2.4 Data analysis

2.4.1 Pre-processing

A standard sequential fMRI processing protocol was carried out using FSL (FMRIB Software Library v6.0, Analysis Group, FMRIB, Oxford, United Kingdom) which included slice scan time correction, temporal high-pass filtering, inter-scan head motion-correction, slight spatial-smoothing, and normalization to standard MNI152 space. Normalization to the standard MNI152 space enables inter-patient comparisons and was used to analyze and interpret the results obtained.

Scans were cropped around each stimulation period starting from the scan immediately before stimulation and covering the stimulation plus 10 s ([-2.5:10s]). Scans were averaged for each condition and hemifield, resulting in 4D matrices of 90 × 90 × 90 voxels × 8 TRs. All data were normalized to have zero mean and unit variance. Each condition and hemifield was trained and tested separately.

2.4.2 Data availability statement

High-quality medical data are difficult to produce, especially when dealing with expensive and time-consuming data sources such as fMRI. Although our study population was larger than similar literature, dataset size is still key when applying DL. As overfitting

mitigation tools, we chose a slim network architecture, limiting trainable parameters, used dropout regularization to simulate an ensemble of NNs with similar yet distinct model configurations, and incorporated artificial data augmentation per batch to increase training variability.

As the name implies, data augmentation is an approach to artificially expand the dataset by generating new samples from existing ones to introduce variability through various manipulations. In this work, we used both standard and specially designed operations. Random translations, rotations, and normal noise addition were used. In addition, we randomly sampled repetitions for averaging, i.e., instead of averaging all 10 repetitions per hemifield per condition, a random combination of between 7 and 9 repetitions was averaged instead (randomly selected from 175 total combinations).

2.4.3 Classification models

CNNs are the gold-standard for image classification (Alzubaidi et al., 2021). This type of network consists of several layers, but is characterized by, and therefore named after, the convolutional layer. This layer uses scanning kernels to convolve the input, which increases its efficiency.

A deep CNN can create more complex abstractions and theoretically solve more complex classification tasks. However, the problem of exploding/vanishing gradients arises. The ResNet architecture, proposed in He et al. (2016), introduced shortcut connections, which mitigate these problems. These shortcut connections are added to the outputs of the stacked layers. ResNet is a modular architecture that stacks multiple residual blocks, i.e., layer groups composed of stacked layers and a shortcut connection that skips them.

The architecture used consists of a 3D convolutional layer followed by four residual blocks and is schematically shown in the [Supplementary materials](#). Batch normalization (BN) and ReLu (Rectified Linear Unit) activation are used to improve stability and introduce non-linearity in the training process, respectively. All convolutions use a stride of one and 3×3 kernels. Within the residual blocks there is a max-pooling layer to halve the feature maps using a 2×2 kernel and stride two. The final classification block includes a fully connected layer, dropout, and a sigmoid activation function to classify each scan.

2.4.4 Training and testing

All scans belonging to DS2 subjects were used for testing only. This separation ensures the independence of testing and the validity of the results. For each model, 10% of DS1 scans were used for hyperparameter tuning. The binary cross-entropy cost-function was used as the loss, which was backpropagated to update the weights of the NN using the stochastic gradient descent optimizer algorithm.

Results on DS2 were obtained using the best performing hyperparameter combination, evaluated independently for each model in the respective tuning set: grid-search to select dropout-rate, learning-rate, momentum, and training steps. [Figure 1](#) summarizes the study workflow.

All models were built, trained, and tested in Python 3.7.9 using the Keras framework with TensorFlow backend. Weights were initially set using a Glorot uniform initializer and biases were set to zero. Network training was performed on a desktop equipped with an AMD

Ryzen 93,900x CPU @3.8 GHz, 64 GB RAM and a Nvidia RTX 3080, using CUDA version 11.5.

2.4.5 Statistical evaluation

To evaluate testing, performance metrics were chosen for the problem at hand and the distribution of the labels in the dataset. The area under the receiver operating characteristic curve (AUC), accuracy, and F1-score were computed. In binary classification tasks these are commonly used to evaluate the performance of a model. Balanced accuracy, sensitivity, and specificity have also been computed.

2.4.6 Interpretability

One of the main drawbacks of non-linear methods like DL is their “black-box” nature. Despite the superior performance shown on many discrete problems, the lack of transparency limits its explainability and hinders its use. Explainable DL is a milestone for clinical translation.

There are several approaches to improve the interpretability of DL methods. Here, we apply deep Taylor decomposition (DTD) (Montavon et al., 2017). The rationale is to decompose the final decision into the individual contributions of each given input by means of relevance backpropagation, i.e., to trace-back contributions from output to input by Taylor decomposition. Ultimately, heatmaps are generated for each classification decision, representing the relative relevance of each voxel to the final classification. Leveraging the individual transformation matrices to the standardized MNI152 space, we transform and average the heatmaps to derive a global response. This approach allows us to identify the brain structures/regions that are more important for classification.

2.4.7 Cluster-based analysis

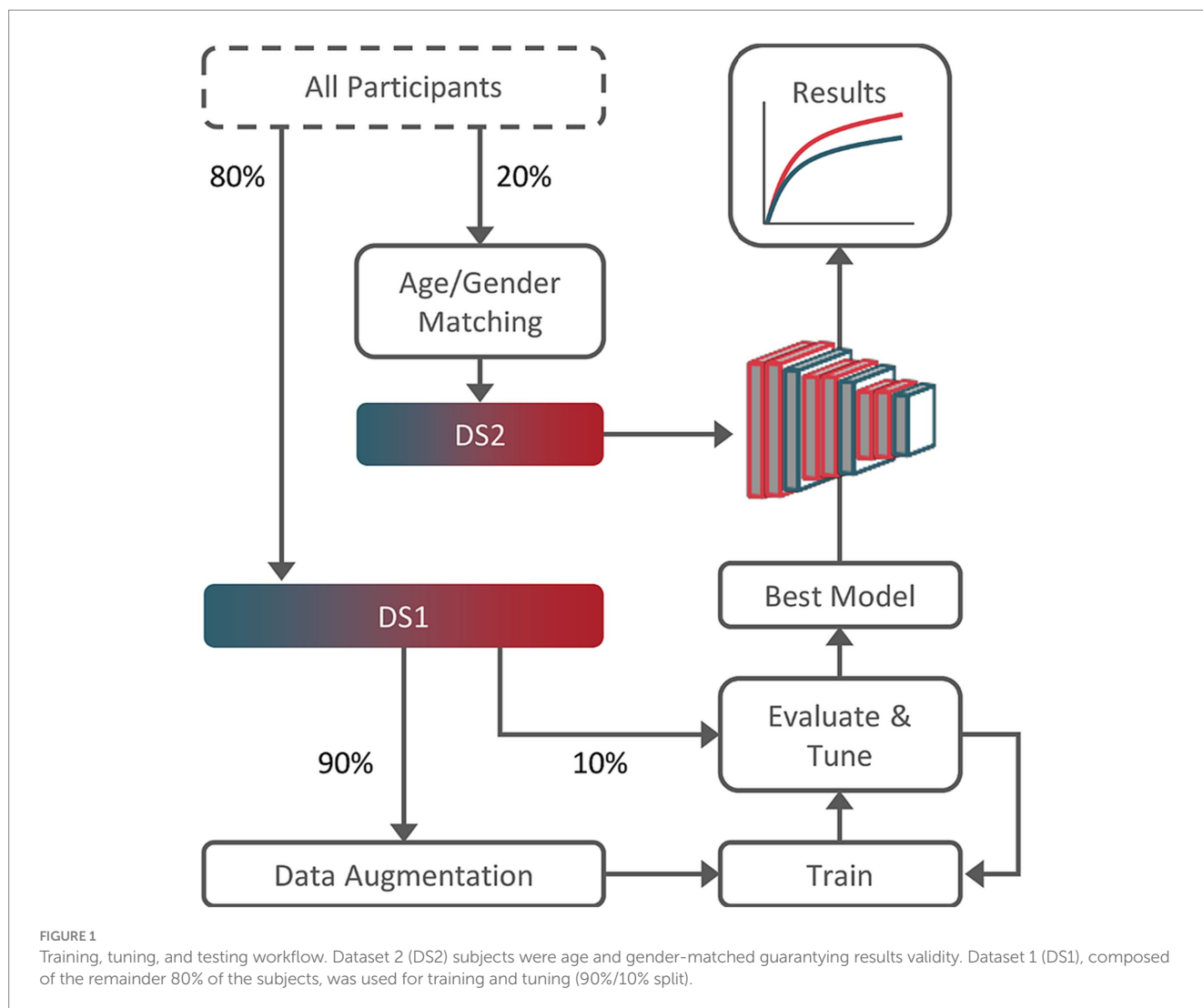
The heatmaps can be further processed to define clusters. We applied an empirically determined threshold to the average heatmap, and defined clusters based on neighborhood connectivity. These were further processed with morphological erosion followed by dilation (morphological open operation). Clusters smaller than 20 voxels were ignored. The resulting relevance clusters can be used as regions of interest (ROIs) for subsequent analysis.

We applied deconvolution analysis within the defined ROIs to separate the contributions of the different events. Instead of one predictor per condition, a fixed number of shifted stick predictors are defined, covering the 20s post-stimulus. It is an alternative to traditional GLM analysis. Its main advantage is its flexibility, since the HRF is not defined *a priori*, but is instead estimated for each event type. This is relevant in the case of diabetes, where standard HRF is not guaranteed.

3 Results

3.1 Data characteristics

A total of 51 T2DM and 67 control subjects were selected. Participants were randomly divided into two datasets: DS1 used for training and tuning of the network, and DS2 used for testing. DS2 is age- and gender-matched, with a mean age of 50.6 ± 1.4 and 51.6 ± 1.8 years for controls and T2DM subjects, respectively. The complete characterization of DS1 and DS2 is available in the



Supplementary materials. All participants underwent three-dimensional anatomical MPRAGE scanning and event-related task-based functional scanning. For the latter, patients were asked to compare the speed of two white dots, a reference and a target dot, and to select the fastest of the two. Overall response accuracy during scanning did not show a significant difference between controls and T2DM subjects (data not shown). This study was conducted under performance-matched conditions with physically matched stimuli.

3.2 fMRI data reveal changes in the brain of patients

Motion-corrected DS1 scans were used to train and tune a CNN to discriminate between T2DM and control scans. The performance was evaluated on DS2 motion-corrected scans. The AUC, accuracy, and F1-score results per condition and hemifield are shown in Figures 2A,B.

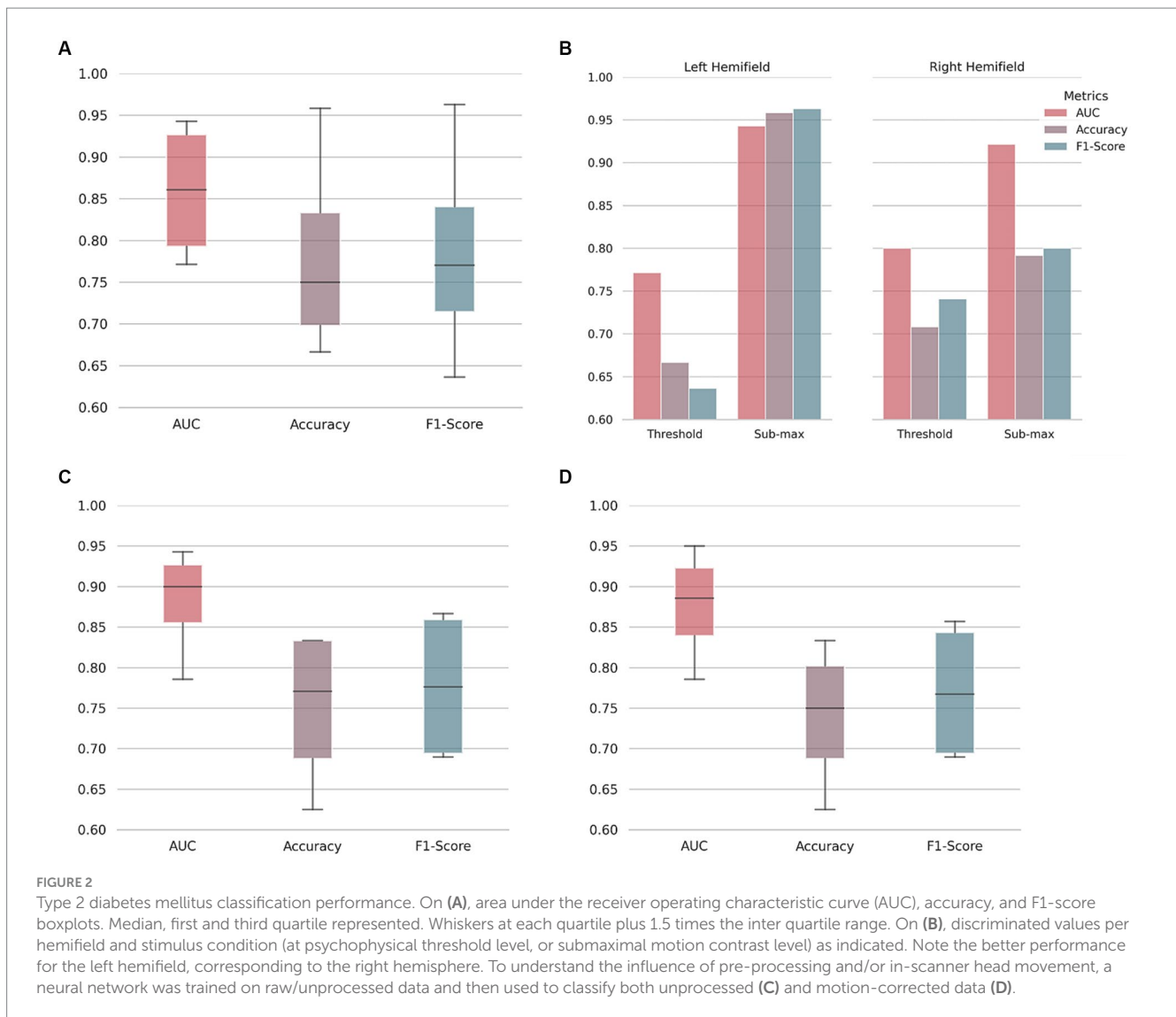
As shown, performance ranges from a minimum of 0.771 AUC and 66.7% accuracy (threshold motion contrast condition, right hemifield) to 0.943 AUC and 95.8% accuracy (sub-maximum motion contrast condition, left hemifield). Discrimination performance for

the sub-maximum condition is better than for the threshold condition in both hemifields, as expected from the larger stimulus signal to noise ratio (serving also as a sanity check). Our results indicate that there are measurable changes in the fMRI signals of T2DM patients compared to controls. Balanced accuracy, sensitivity, and specificity results are shown in the Supplementary materials.

3.3 Head movement does not affect performance

DL is a broad-spectrum classifier. With NNs, layer by layer increasingly more complex representations are created, and, theoretically, no feature representation is off-limits. Despite the exclusion criteria used (subjects with in-scanner head movement greater than 3 mm were excluded), it is important to ensure that neither the head movement in the scanner, nor the pre-processing typically applied to correct for it, is partially responsible for the discrimination performance.

To understand the influence of pre-processing and/or in-scanner head movement, data were compared at two processing stages: unprocessed (or raw) and motion-corrected. A CNN was trained and



tuned with raw fMRI data. This network was then used to classify both raw and motion-corrected data (Figures 2C,D). As shown, although there is some variability, similar test performance was achieved for the two processing stages. Results suggest that our classification approach has no specific bias due to pre-processing or in-scanner head movement up to the tested limit of 3 mm.

3.4 Backpropagation shows relevant brain regions

To shed light on the decision process, we derived the individual pixel contributions using DTD (Montavon et al., 2017) and generated relevance heatmaps (Figure 3).

The resulting heatmaps were then transformed into standardized MNI152 space to be averaged across subjects. Average relevance patterns calculated independently for controls and T2DM subjects showed similar patterns (see Supplementary materials). From the resulting average heatmaps, we could discern six semi-consistent agglomerations across the four models (two conditions per hemifield). A global average relevance heatmap was computed and clusters were

determined (Figures 3B–G). As shown, clusters are located approximately in the superior frontal gyrus, angular gyrus, visual cortex, insula, cerebellum, and thalamus. The sagittal, coronal, and transverse planes are shown for each cluster, centered on the peak relevance voxel. The location of the peak relevance voxel (MNI152 space) and volume for each cluster are shown in the Supplementary materials.

3.5 Cluster-based hemodynamic response function

We resorted to deconvolution-based GLM analysis to more accurately represent the time course of the neural activity estimated from the measured BOLD response. In Figure 4, we show the global average of the estimated HRF for each group (controls and T2DM subjects), computed for both threshold and sub-maximum events (in terms of speed contrast) and over the six derived clusters. In the specific case of the deconvolution GLM, the beta weights of the GLM allow the reconstruction of the entire BOLD response.

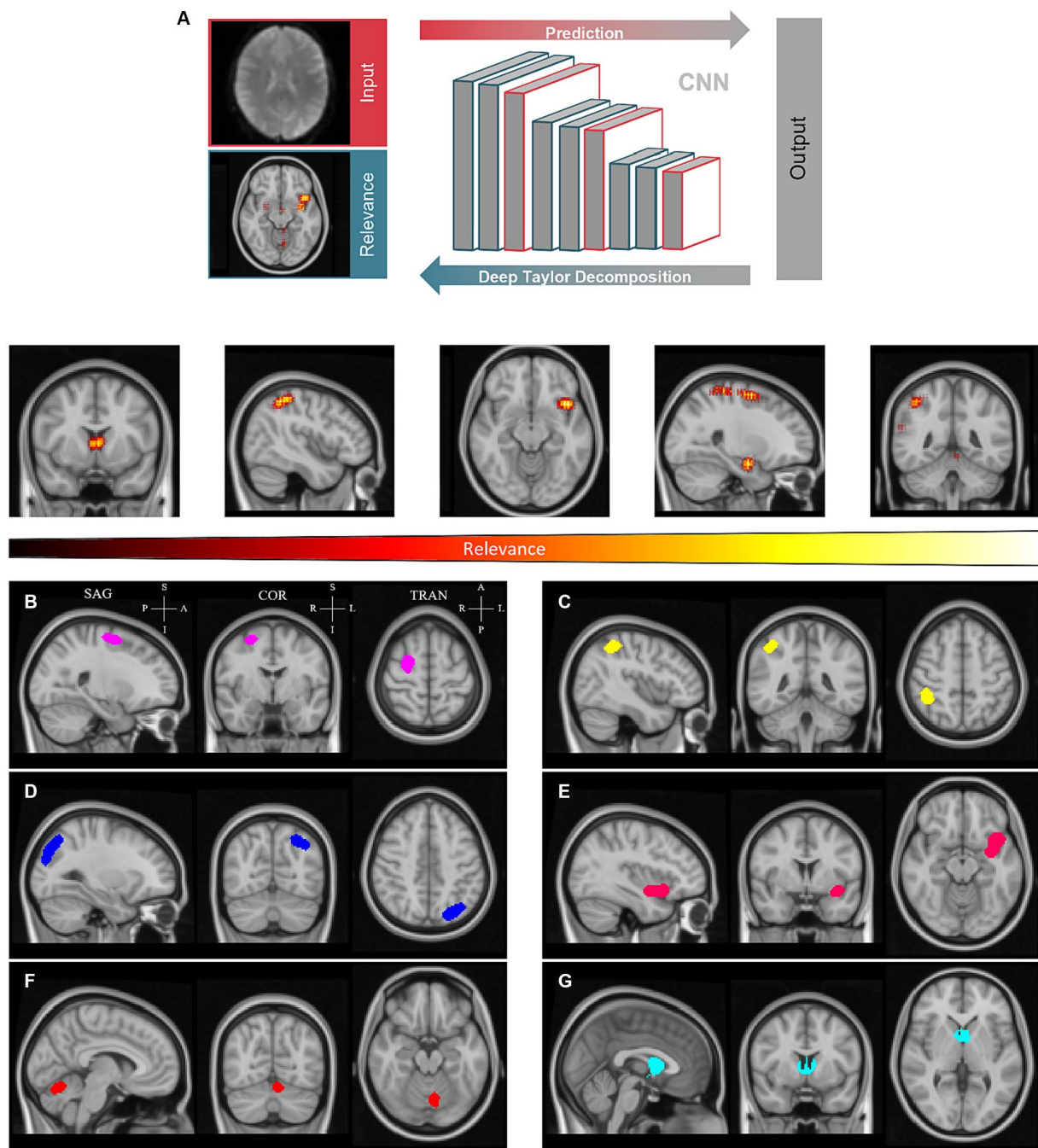


FIGURE 3 Clustering of relevance heatmaps computed by deep Taylor decomposition to define regions of interest. **(A)** - Individual pixel contributions were backpropagated over the convolutional neural network (CNN) to generate relevance heatmaps. **(B–G)** - Averaged heatmaps were thresholded and labeled to create relevance clusters. Sagittal (SAG), coronal (COR), and transverse (TRA) planes are shown for each cluster. These are located approximately in the superior frontal gyrus **(B)**, angular gyrus **(C)**, extrastriate visual and parietal cortex **(D)**, insula **(E)**, cerebellum **(F)**, and thalamus **(G)**, left to right and, top to bottom, respectively. Clusters are represented with distinct colors for the sake of convenience. P – Posterior; A – Anterior; S – Superior; I – Inferior.

In T2DM patients, the characteristic undershoot of the HRF before returning to baseline seems to be less pronounced compared to controls (Figures 4A–D). The HRF peaks lower in T2DM subjects in most ROIs. However, the ROI located in the superior frontal gyrus appears to be an exception (Figures 4A–D). Although the general shape is similar between controls and T2DM subjects, the HRF of the latter appears to be delayed compared to controls. This effect, present in several ROIs, is most pronounced in Figures 4G,H.

4 Discussion

For the first time, we have combined event-related task-based fMRI imaging, deconvolution GLM analysis, and novel DL methods. This innovative methodology offers a unique perspective to validate in a data-driven manner changes in the BOLD response of T2DM patients. Neither DL nor deconvolution GLM use any *a priori* information. With this fully data-driven approach, we unbiasedly

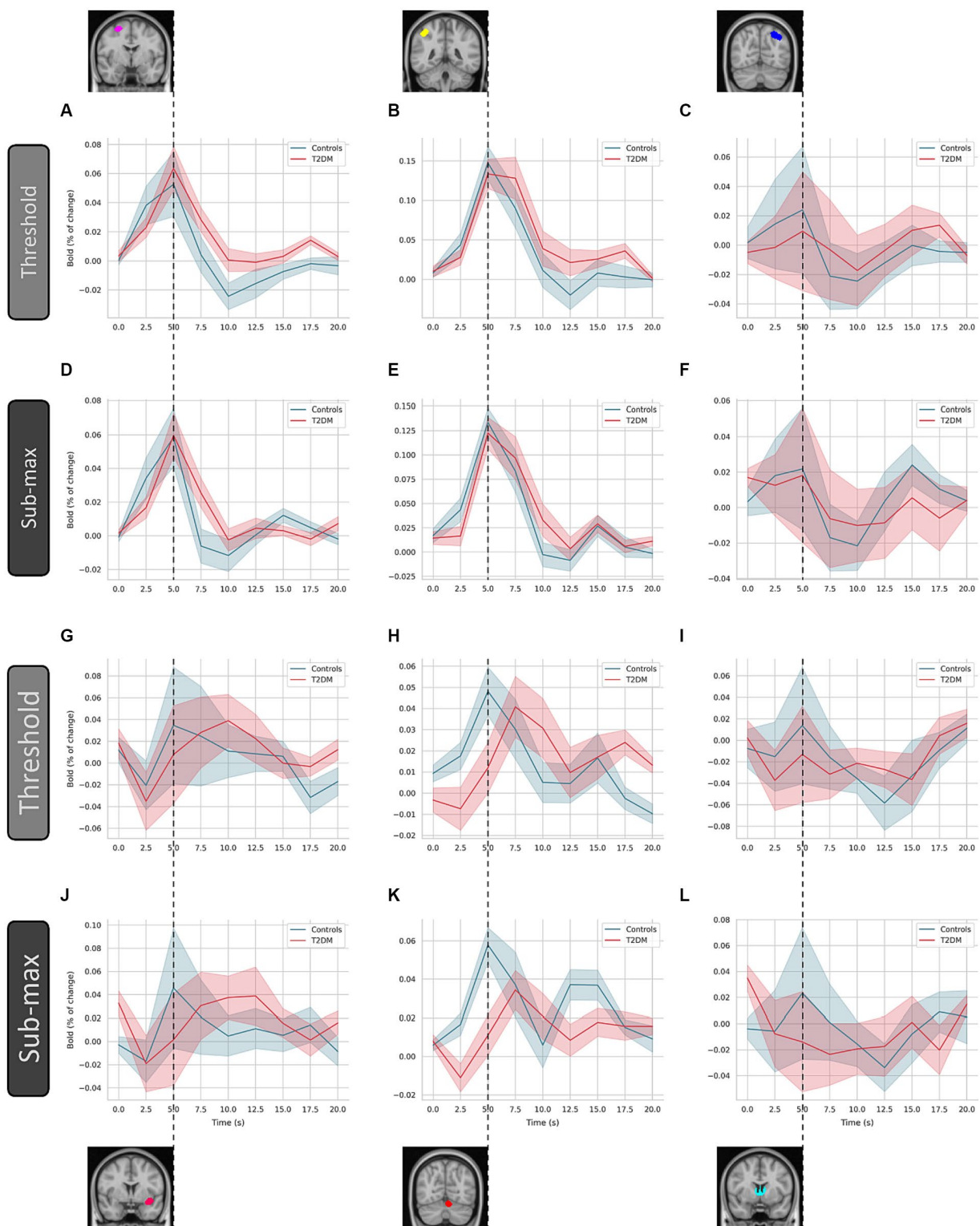


FIGURE 4
 Estimated hemodynamic response function by deconvolution-based analysis. The hemodynamic response function was computed over the six indicated clusters for controls and type 2 diabetes patients (T2DM). These are (left to right, top to bottom) the superior frontal gyrus (A and D), angular gyrus (B and E), visual cortex (C and F), insula (G and J), cerebellum (H and K), and thalamus (I and L), respectively. Average and standard deviation are represented (line and shaded area, respectively), for each condition (threshold and sub-maximum as indicated). Dashed line highlights the time of peak response of control subjects.

demonstrated that T2DM patients have an altered hemodynamic event-related response compared to age- and gender-matched controls, which may therefore represent a potential disease biomarker.

We used DL to model normal and T2DM task-based (visual speed discrimination) fMRI scans. The obtained results show that discrimination between the two groups was possible and accurate.

The model trained on the sub-maximum condition (the condition with the highest speed contrast) presented on the left hemifield accurately predicted over 95% of the scans. This is reassuring since optimal classification is attained using the best stimulus contrast, and at the hemisphere that dominates for spatial vision, thereby confirming biological plausibility. For the threshold condition (lowest speed contrast, left hemifield), the trained model had an accuracy rate of 66.7%. These hemifield differences are physiologically meaningful and might be explained by known hemispheric differences in the cortical spatial processing (Silva et al., 2008). The left visual hemifield corresponds to the brain's right hemifield, which dominates in cortical visuospatial processing (Silva et al., 2008). Performance variability may also reflect task complexity.

To investigate the influence of both the head motion in the scanner and/or the preprocessing applied to correct it, a CNN was trained and tuned on raw/unprocessed fMRI data and then used to classify both unprocessed and motion-corrected data. The performance of the models was similar in both cases. The performance of the network trained on raw data and the network trained on motion-corrected data was also comparable. These results suggest that the influence is negligible.

The relevance clusters obtained from the DTD heatmaps were used as ROIs for subsequent deconvolution analysis to understand the differences between T2DM patients and control subjects. T2DM patients showed an altered hemodynamic response. The results of this data-driven approach are consistent with our hypothesis proposed in Duarte et al. (2015, 2023). In response to increased neuronal activation, there is an increase in oxygen demand and, consequently, an increase in blood supply. In diseased subjects, this autoregulatory system may be disrupted, resulting in alterations in the standard BOLD signal and HRF. Physiological changes in the brain vasculature, prior to any evidence of over structural lesions, may affect the normal response. The demonstrated HRF differences challenges the conventional analysis of task-based fMRI in T2DM patients with the standard HRF and its appropriateness in neurocognitive paradigms. These altered HRF patterns are promising biomarkers that may lead to a better understanding of the neurophysiological changes in T2DM. Overall, we demonstrated the potential of DL to provide novel insight into neuropathology and highlighted the need for interpretable DL methods in healthcare research.

4.1 Strengths and limitations

MR structural and functional imaging are indispensable tools in the research, diagnosis, and follow-up of neuropathology. However, MR remains a time- and labor-intensive data source which inherently limits data availability. Task-based protocols increase specificity but intrinsically restrict inter-study sharing. One of the reasons why the largest datasets rely on resting state. However, resting state cannot directly capture stimulus-related functional properties and event-related responses. These are critical for unraveling HRFs. Given these limitations, our study had a larger sample size than similar studies. However, it is important to note that a more extensive sample size would further strengthen the study's findings. This study serves as an

initial step, highlighting the need for future research and to encompass larger cohorts. Mainly, the integration of multicentric datasets covering the entire spectrum of T2DM will facilitate the development of more robust methodologies and conclusions.

Because NNs are powerful but indiscriminate learners, controlling bias is crucial. We aimed to maximize the participant pool, while ensuring the validity of the results. DS2, for testing, was age and gender-matched, while DS1, for training and tuning, was not, which is less critical, but may impact performance. The performance achieved is valid but could ultimately be improved further.

Despite the performance advantage of DL against other machine learning methods, the lack of transparency inherent to non-linear methods means that one cannot know what factors its decision is being made upon. The biological plausibility of classification performance is addressed above. In this work, we applied DTD to derive heatmaps representing the relative importance of each voxel for the classification between controls and T2DM. Functional scans were co-registered with structural MRI scans and, in turn, normalized to the MNI152 space. This allowed us to transform the resulting heatmaps to standardized space for group analysis. Inter-subject averaging revealed relevance clusters located in the superior frontal and angular gyrus (Zhou et al., 2010; García-Casares et al., 2014; Tan et al., 2019; Qi et al., 2021), visual extrastriate cortex (Newsome and Pare, 1988; Yao et al., 2021), insula (Chang et al., 2006; Xia et al., 2017), cerebellum (Chen et al., 2015; Yao et al., 2021), and thalamus (Chen et al., 2015), all regions related to the disease and/or task (speed discrimination processing and perceptual decision). It is revealing that these regions were of more importance in this data-driven approach, as these are consistent with previous observations in T2DM, which helps to substantiate the classification obtained.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors upon formal and reasonable request.

Ethics statement

The studies involving humans were approved by Ethics Committee of the Faculty of Medicine of the University of Coimbra. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

PG: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft. PS: Formal analysis, Methodology, Visualization, Writing – review & editing. JD: Data curation, Formal analysis, Writing – review & editing. JC: Data curation, Formal analysis, Writing – review & editing. CM: Data curation, Resources, Writing – review & editing. LG: Data curation,

Resources, Writing – review & editing. RB: Formal analysis, Methodology, Supervision, Writing – review & editing. MC-B: Conceptualization, Formal analysis, Funding acquisition, Methodology, Project administration, Resources, Supervision, Validation, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. FCT DSAIPA/DS/0041/2020, FCT/UIDB/4950/2020 and FCT/UIDP/4950/2020 and EASD Innovative Outcomes 2019.

Acknowledgments

We thank all the participants involved. We also thank Carlos Ferreira and João Marques from the Institute for Nuclear Sciences Applied to Health from the University of Coimbra, for their help with the MRI acquisitions.

References

- Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., et al. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *J. Big Data* 8:53. doi: 10.1186/s40537-021-00444-8
- Bancks, M. P., Alonso, A., Gottesman, R. F., Mosley, T. H., Selvin, E., and Pankow, J. S. (2017). Brain function and structure and risk for incident diabetes: the atherosclerosis risk in communities study. *Alzheimers Dement.* 13, 1345–1354. doi: 10.1016/j.jalz.2017.04.006
- Biessels, G. J., and Despa, F. (2018). Cognitive decline and dementia in diabetes mellitus: mechanisms and clinical implications. *Nat. Rev. Endocrinol.* 14, 591–604. doi: 10.1038/s41574-018-0048-7
- Biessels, G. J., and Reijmer, Y. D. (2014). Brain changes underlying cognitive dysfunction in diabetes: what can we learn from MRI? *Diabetes* 63, 2244–2252. doi: 10.2337/db14-0348
- Biessels, G. J., Strachan, M. W. J., Visseren, F. L. J., Kappelle, L. J., and Whitmer, R. A. (2014). Dementia and cognitive decline in type 2 diabetes and prediabetic stages: towards targeted interventions. *Lancet Diabetes Endocrinol.* 2, 246–255. doi: 10.1016/S2213-8587(13)70088-3
- Chang, P.-Y., Carrel, H., Huang, J.-S., Wang, I.-J., Hou, Y.-C., Chen, W.-L., et al. (2006). Decreased density of corneal basal epithelium and subbasal corneal nerve bundle changes in patients with diabetic retinopathy. *Am J. Ophthalmol.* 142, 488–490.e1. doi: 10.1016/j.ajo.2006.04.033
- Chen, Z., Sun, J., Yang, Y., Lou, X., Wang, Y., Wang, Y., et al. (2015). Cortical thinning in type 2 diabetes mellitus and recovering effects of insulin therapy. *J. Clin. Neurosci.* 22, 275–279. doi: 10.1016/j.jocn.2014.07.014
- Chen, Y.-C., Xia, W., Qian, C., Ding, J., Ju, S., and Teng, G.-J. (2015). Thalamic resting-state functional connectivity: disruption in patients with type 2 diabetes. *Metab. Brain Dis.* 30, 1227–1236. doi: 10.1007/s11011-015-9700-2
- Chhatwal, J. P., and Sperling, R. A. (2012). Functional MRI of mnemonic networks across the Spectrum of Normal aging, mild cognitive impairment, and Alzheimer's disease. *J. Alzheimers Dis.* 31, S155–S167. doi: 10.3233/JAD-2012-120730
- Crisóstomo, J., Duarte, J. V., Moreno, C., Gomes, L., and Castelo-Branco, M. (2021). A novel morphometric signature of brain alterations in type 2 diabetes: patterns of changed cortical gyrification. *Eur. J. Neurosci.* 54, 6322–6333. doi: 10.1111/ejn.15424
- Cui, Y., Jiao, Y., Chen, Y.-C., Wang, K., Gao, B., Wen, S., et al. (2014). Altered spontaneous brain activity in type 2 diabetes: a resting-state functional MRI study. *Diabetes* 63, 749–760. doi: 10.2337/db13-0519
- Duarte, J. V., Guerra, C., Moreno, C., Gomes, L., and Castelo-Branco, M. (2023). Changes in hemodynamic response function components reveal specific changes in neurovascular coupling in type 2 diabetes. *Front. Physiol.* 13:13. doi: 10.3389/fphys.2022.1101470
- Duarte, J. V., Pereira, J. M., Quendera, B., Raimundo, M., Moreno, C., Gomes, L., et al. (2015). Early disrupted neurovascular coupling and changed event level hemodynamic response function in type 2 diabetes: an fMRI study. *J. Cereb. Blood Flow Metab.* 35, 1671–1680. doi: 10.1038/jcbfm.2015.106

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fninf.2023.1321178/full#supplementary-material>

- García-Casares, N., Berthier, M. L., Jorge, R. E., Gonzalez-Alegre, P., Gutiérrez Cardo, A., Rioja Villodres, J., et al. (2014). Structural and functional brain changes in middle-aged type 2 diabetic patients: a cross-sectional study. *J. Alzheimers Dis.* 40, 375–386. doi: 10.3233/JAD-131736
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). “Deep residual learning for image recognition,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. (Las Vegas, USA), 770–778.
- Horikawa, T., and Kamitani, Y. (2017). Generic decoding of seen and imagined objects using hierarchical visual features. *Nat. Commun.* 8:15037. doi: 10.1038/ncomms15037
- International Diabetes Federation. (2021). *IDF Diabetes Atlas, 10th edn*. Brussels, Belgium. Available at: <https://www.diabetesatlas.org>
- Kim, J., Calhoun, V. D., Shim, E., and Lee, J.-H. (2016). Deep neural network with weight sparsity control and pre-training extracts hierarchical features and enhances classification performance: evidence from whole-brain resting-state functional connectivity patterns of schizophrenia. *NeuroImage* 124, 127–146. doi: 10.1016/j.neuroimage.2015.05.018
- Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., et al. (2017). A survey on deep learning in medical image analysis. *Med. Image Anal.* 42, 60–88. doi: 10.1016/j.media.2017.07.005
- Liu, D., Duan, S., Zhou, C., Wei, P., Chen, L., Yin, X., et al. (2018). Altered brain functional hubs and connectivity in type 2 diabetes mellitus patients: a resting-state fMRI study. *Front. Aging Neurosci.* 10:10. doi: 10.3389/fnagi.2018.00055
- Liu, H., Liu, J., Peng, L., Feng, Z., Cao, L., Liu, H., et al. (2019). Changes in default mode network connectivity in different glucose metabolism status and diabetes duration. *NeuroImage Clin.* 21:101629. doi: 10.1016/j.nicl.2018.101629
- Moheet, A., Mangia, S., and Seaquist, E. R. (2015). Impact of diabetes on cognitive function and brain structure. *Ann. N. Y. Acad. Sci.* 1353, 60–71. doi: 10.1111/nyas.12807
- Montavon, G., Lapuschkin, S., Binder, A., Samek, W., and Müller, K.-R. (2017). Explaining nonlinear classification decisions with deep Taylor decomposition. *Pattern Recogn.* 65, 211–222. doi: 10.1016/j.patcog.2016.11.008
- Moran, C., Phan, T. G., Chen, J., Blizzard, L., Beare, R., Venn, A., et al. (2013). Brain atrophy in type 2 diabetes. *Diabetes Care* 36, 4036–4042. doi: 10.2337/dc13-0143
- Newsome, W., and Pare, E. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *J. Neurosci.* 8, 2201–2211. doi: 10.1523/JNEUROSCI.08-06-02201.1988
- Qi, F., Zhang, D., Gao, J., Tang, M., Wang, M., Su, Y., et al. (2021). Functional disconnection of the angular gyrus related to cognitive impairment in patients with type 2 diabetes mellitus. *Front. Hum. Neurosci.* 15:15. doi: 10.3389/fnhum.2021.621080
- Ryan, J. P., Fine, D. E., and Rosano, C. (2014). Type 2 diabetes and cognitive impairment. *J. Geriatr. Psychiatry Neurol.* 27, 47–55. doi: 10.1177/0891988713516543
- Santana, A. N., Cifre, I., de Santana, C. N., and Montoya, P. (2019). Using deep learning and resting-state fMRI to classify chronic pain conditions. *Front. Neurosci.* 13:13. doi: 10.3389/fnins.2019.01313

- Sarraf, S., and Tofghi, G. Deep learning-based pipeline to recognize Alzheimer's disease using fMRI data (2016). *Futur. Technol. Conf. IEEE 2016*, 816–820. doi: 10.1109/FTC.2016.7821697
- Silva, M. F., Maia-Lopes, S., Mateus, C., Guerreiro, M., Sampaio, J., Faria, P., et al. (2008). Retinal and cortical patterns of spatial anisotropy in contrast sensitivity tasks. *Vis. Res.* 48, 127–135. doi: 10.1016/j.visres.2007.10.018
- Suk, H.-I., Wee, C.-Y., Lee, S.-W., and Shen, D. (2016). State-space model with deep learning for functional dynamics estimation in resting-state fMRI. *NeuroImage* 129, 292–307. doi: 10.1016/j.neuroimage.2016.01.005
- Tan, X., Liang, Y., Zeng, H., Qin, C., Li, Y., Yang, J., et al. (2019). Altered functional connectivity of the posterior cingulate cortex in type 2 diabetes with cognitive impairment. *Brain Imaging Behav.* 13, 1699–1707. doi: 10.1007/s11682-018-0017-8
- Tiehuis, A. M., Vincken, K. L., van den Berg, E., Hendrikse, J., Manschot, S. M., Mali, W. P. T. M., et al. (2008). Cerebral perfusion in relation to cognitive function and type 2 diabetes. *Diabetologia* 51, 1321–1326. doi: 10.1007/s00125-008-1041-9
- Warren, S. L., and Moustafa, A. A. (2023). Functional magnetic resonance imaging, deep learning, and Alzheimer's disease: a systematic review. *J. Neuroimaging* 33, 5–18. doi: 10.1111/jon.13063
- Xia, W., Chen, Y.-C., and Ma, J. (2017). Resting-state brain anomalies in type 2 diabetes: a Meta-analysis. *Front. Aging Neurosci.* 9:9. doi: 10.3389/fnagi.2017.00014
- Xia, W., Wang, S., Sun, Z., Bai, F., Zhou, Y., Yang, Y., et al. (2013). Altered baseline brain activity in type 2 diabetes: a resting-state fMRI study. *Psychoneuroendocrinology* 38, 2493–2501. doi: 10.1016/j.psyneuen.2013.05.012
- Xiong, Y., Chen, X., Zhao, X., Fan, Y., Zhang, Q., and Zhu, W. (2020). Altered regional homogeneity and functional brain networks in type 2 diabetes with and without mild cognitive impairment. *Sci. Rep.* 10:21254. doi: 10.1038/s41598-020-76495-3
- Yao, L., Yang, C., Zhang, W., Li, S., Li, Q., Chen, L., et al. (2021). A multimodal meta-analysis of regional structural and functional brain alterations in type 2 diabetes. *Front. Neuroendocrinol.* 62:100915. doi: 10.1016/j.yfrne.2021.100915
- Zaharchuk, G., Gong, E., Wintermark, M., Rubin, D., and Langlotz, C. P. (2018). Deep learning in neuroradiology. *Am. J. Neuroradiol.* 39, 1776–1784. doi: 10.3174/ajnr.A5543
- Zhang, T., Shaw, M., and Cherbuin, N. (2022). Association between type 2 diabetes mellitus and brain atrophy: a Meta-analysis. *Diabetes Metab. J.* 46, 781–802. doi: 10.4093/dmj.2021.0189
- Zhou, H., Lu, W., Shi, Y., Bai, F., Chang, J., Yuan, Y., et al. (2010). Impairments in cognition and resting-state connectivity of the hippocampus in elderly subjects with type 2 diabetes. *Neurosci. Lett.* 473, 5–10. doi: 10.1016/j.neulet.2009.12.057