NeuralGait: Assessing Brain Health Using Your Smartphone

HUINING LI*, University at Buffalo, USA
HUAN CHEN*, Case Western Reserve University, USA
CHENHAN XU, University at Buffalo, USA
ZHENGXIONG LI, University of Colorado Denver, USA
HANBIN ZHANG, University at Buffalo, USA
XIAOYE QIAN, Case Western Reserve University, USA
DONGMEI LI, University of Rochester Medical Center, USA
MING-CHUN HUANG, Duke Kunshan University, China and Suzhou Huanmu Intelligence Technology Co., Ltd., China
WENYAO XU, University at Buffalo, USA

Brain health attracts more recent attention as the population ages. Smartphone-based gait sensing and analysis can help identify the risks of brain diseases in daily life for prevention. Existing gait analysis approaches mainly hand-craft temporal gait features or developing CNN-based feature extractors, but they are either prone to lose some inconspicuous pathological information or are only dedicated to a single brain disease screening. We discover that the relationship between gait segments can be used as a principle and generic indicator to quantify multiple pathological patterns. In this paper, we propose *NeuralGait*, a pervasive smartphone-cloud system that passively captures and analyzes principle gait segments relationship for brain health assessment. On the smartphone end, inertial gait data are collected while putting the smartphone in the pants pocket. We then craft local temporal-frequent gait domain features and develop a self-attention-based gait segment relationship encoder. Afterward, the domain features and relation features are fed to a scalable RiskNet in the cloud for brain health assessment. We also design a pathological hot update protocol to efficiently add new brain diseases in the RiskNet. *NeuralGait* is practical as it provides brain health assessment with no burden in daily life. In the experiment, we recruit 988 healthy people and 417 patients with a single or combination of PD, TBI, and stroke, and evaluate the brain health assessment using a set of specifically designed metrics including global accuracy, exact accuracy, sensitivity, and false alarm rate. We also demonstrate the generalization (e.g., analysis of feature effectiveness and model efficiency) and inclusiveness of *NeuralGait*.

CCS Concepts: • **Human-centered computing → Ubiquitous and mobile computing systems and tools**; • **Applied computing → Life and medical sciences**.

Additional Key Words and Phrases: Multi-label brain diseases, gait analysis, smartphone

*Both authors contributed equally to the paper

Address comments to wenyaoxu@buffalo.edu.
Authors' address: Huining Li; Chenhan Xu; Hanbin Zhang; Wenyao Xu, University at Buffalo, the State University of New York, Department of Computer Science and Engineering, Amherst, NY, 14260, USA. Huan Chen; Xiaoye Qian, Case Western Reserve University, Electrical and Computer Engineering Department, Cleveland, OH, 44106, USA. Zhengxiong Li, University of Colorado Denver, Department of Computer Science and Engineering, Denver, CO, 80204, USA. Dongmei Li, University of Rochester Medical Center, Rochester, NY, 14642, USA. Ming-chun Huang, Duke Kunshan University, Suzhou, Jiangsu, 215316, China.

2474-9567/2022/12-ART169 $15.00
https://doi.org/10.1145/3569476

ACM Reference Format:
Huining Li, Huan Chen, Chenhan Xu, Zhengxiong Li, Hanbin Zhang, Xiaoye Qian, Dongmei Li, Ming-Chun Huang, and Wenyao Xu. 2022. NeuralGait: Assessing Brain Health Using Your Smartphone. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 4, Article 169 (December 2022), 28 pages. https://doi.org/10.1145/3569476

1 INTRODUCTION

The human brain is the command center for the nervous system and controls behaviors, cognition, and emotion using a complex function [1]. As the population ages, the burden of neurological disorders and challenges for preserving brain health encounter increase. Till now, more than 10 million people live with Parkinson's disease (PD) [2], and 55 million people live with dementia worldwide [3]. In each year of the United States, 795,000 Americans have a stroke [4], and more than 223,000 traumatic brain injury (TBI)-related hospitalizations [5]. In-time diagnosis and treatment of brain diseases are critical to slow or reverse the disease progression for pursuing health and longevity. However, individuals are seldom aware of the early signs and progression of brain diseases and thereby will not seek clinical diagnosis and treatment. As clinical researchers validated the association between neural pathology and gait impairment [6, 7], gait is considered as an effective behavioral biomarker of brain diseases. With the advancement of micro-electro-mechanical systems (MEMS) technologies [8], the smartphone's built-in inertial sensors can provide always-on monitoring of gait without the restriction of locations and environmental conditions. Therefore, smartphone-based gait analysis has been a primary trend to screen brain diseases and aid the cue of the onset risks in daily life.

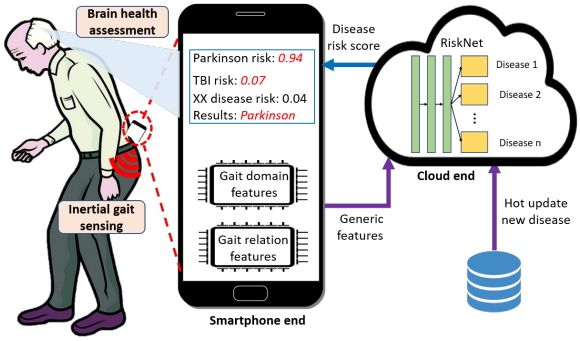


Fig. 1. NeuralGait passively collects a user's inertial gait data with the smartphone in daily life for risk estimation of multi-label brain diseases.

Existing smartphone-based gait analysis approaches mainly have two directions: 1) *Temporal gait features analysis*, such as swing/stance time, double/single support time, stride regularity, and freeze index [9]. Although these features can capture explainable gait properties, they are hand-crafted and prone to lose some inconspicuous but key pathological information that compromises brain disease screening accuracy. 2) *Data-driven gait feature learning*, i.e., developing CNN-based models (e.g., ResNet) [10, 11] or autoencoder-based models as feature extractors [12]. However, the CNN-based feature extractor needs to be specifically trained for a certain neural disease screening because its nature is a data-driven non-linear 2D filter that focuses on learning local pathological information of given gait data. The autoencoder-based model is not dedicated to a single brain disease, but the learned features cannot be explained and quantified. To summarize, although gait is considered an effective behavioral biomarker for screening a broad spectrum of brain diseases, high-capacity pathological information hidden in gait is largely unexploited and unexplained.

It is a fact that gait is generated by the complex coordination of the nervous system and musculoskeletal system [13]. Based on action potentials and chemical neurotransmitters, the brain generates stepping commands and disseminates them through neurons till it excites the muscle fibers and initiates walking. During walking, the behaviors of heel-strike and toe-off alternately happen to regulate the swing phase and stance phase, meanwhile, the brain further adjusts the gait based on the sensory feedback for adapting to the environment. On the basis of the gait theory, *we envision that gait segments in a cycle are affected by each other*. If a brain disease changes the gait properties in a certain segment of a gait cycle, it will influence its relationship with other gait segments. Therefore, the relationship between different gait segments can be a **principle** cue for identifying pathological patterns. With the advancement of natural language processing (NLP), the relationship learning among words for understanding the text and speech has made great progress [14]. Inspired by these works, a principle and pervasive smartphone-based gait analysis system can be developed to learn gait segment relationships for multiple brain disease screening.

To realize our system, three challenges need to be addressed: 1) How to design a gait embedding approach (like word embedding in NLP) that preserves the local domain knowledge of gait cycles? 2) How to develop a principle gait segment relationship encoder and quantify the relation features across different brain diseases? 3) In the real practice of brain health assessment, healthy subjects, and diseased subjects are mutually exclusive, but different types of neural diseases are not mutually exclusive, how to develop a general and scalable neural disease detection framework that can identify healthy subjects and screen multi-label neural diseases simultaneously?

In this paper, we propose an end-to-end smartphone-cloud framework, *NeuralGait*, to capture users' principle gait segment relationship from daily life using an in-pocket smartphone for pervasive brain health assessment, as shown in Fig. 1. Specifically, at the smartphone end, we first leverage the built-in sensors, i.e., accelerometer and gyroscope to collect and extract gait cycles. To retain the local domain knowledge of gait, we transform the time-series gait cycle into a spectrogram and split it into multiple patches. Each patch is further embedded with appearance information (i.e., local power distribution) and geometric information (i.e., corresponding frequency component and time slot). After that, we develop a self-attention-based gait segments relationship learning network with these patch embedding as input to extract time-relation features, frequency-relation features, and global relation-aware features. To maximize user experience and save battery, the self-attention model is pruned based on important scores. At the cloud-end, the extracted features are then fed to a scalable RiskNet that consists of a shared module for features refinement, and multiple customized modules (i.e., branches) with each for predicting a type of neural disease. Healthy subjects are identified if the outputs of all branches are negative. We also develop a pathological hot-update protocol to efficiently update new brain disease in the RiskNet, which means weights in the shared module are fixed and only the customized disease prediction branch is updated. Finally, more reliable results of brain health assessment are transmitted to users with the help of our continuous detection protocol.

We evaluate *NeuralGait* on 988 healthy people and 417 patients with a single or combination of PD, TBI, and stroke. In our setting, stroke is a newly introduced disease to the system and the stroke prediction is trained on the customized model. All the experiments are performed on unseen users to guarantee applicability in the real world. We investigate a set of metrics to evaluate the system performance in the aspect of overall multi-label multi-class classification, distinguishing healthy and diseased, and screening each type of neural disease. Our system achieves 80.1% global accuracy for multi-label multi-class disease classification. The false alarm rate for screening each type of neural disease is less than 10%, and the sensitivity for screening stroke reaches up to 76.1%. We further conduct a generalization study to examine the effectiveness and efficiency of our proposed gait analysis approach for unseen neural disease screening.

The contribution of our work is three-fold:

- We perform the first study to identify that the relationship between different gait segments can be used as a principle and effective cue to screen multi-label brain diseases in daily life.
- We develop *NeuralGait*, an end-to-end pervasive system to facilitate brain health assessment. We first capture gait data with in-pocket smartphones, and then craft local temporal-frequent gait domain knowledge and extract gait segment relationship features. Afterward, these principle features are input to a scalable RiskNet in the cloud for multi-label multi-class disease screening. Further, a pathological hot-update protocol is designed to instantly update new diseases in the cloud.
- We evaluate *NeuralGait* on healthy subjects and patients with multi-label diseases. Our system can achieve 80.1% global accuracy and 77.7% exact accuracy for multi-label multi-class disease classification. We further examine the inclusiveness of *NeuralGait* in screening each type of neural disease.

2 BACKGROUND AND PRELIMINARIES

2.1 Gait Rationale

Gait is a complex task that involves the coordination of the nervous system and musculoskeletal system.[15]. Central pattern generators in the spinal cord produce rhythmic stepping [16]. To adapt to the changes in the environment, the subject needs to regulate the gait based on the sensory feedback during the walk. Specifically, the vestibular cortex provides the position and motion information of the head as the person walks. The visual cortex provides the size and location information when approaching an obstacle. These are integrated and transmitted to the supplementary motor area and the premotor area of the cerebral cortex to create a motor program for intentional postural adjustments and limb motion. The cerebellum also receives sensory information and compares the actual and intended stepping pattern. If there is a discrepancy, the cerebellum decides the appropriate correction and relays the information to the brainstem and motor cortex for coordination [17].

2.2 Fundamental Brain Diseases and Gait Segment Relationship

Based on the pathogenesis mechanisms, major brain disorders can be divided into two types: neurodegeneration (e.g., Parkinson's disease, dementia) and structural damage (e.g., traumatic brain injury, stroke, transient ischemic attack). The difference in pathogenesis mechanism causes various visible pathological gait properties. In essence, the principle manifestation is the changed relationship between different gait segments.

Specifically, neurodegeneration is the progressive loss of structure or function of neurons and may involve the shrinkage of the cerebral cortex and hippocampus [18]. For example, Parkinson's disease is characterized by the death of dopaminergic neurons in the substantia nigra areas, which will impede the signals controlling the motion be transmitted to proper body regions. It can induce three gait abnormality phenotypes, i.e., freezing of gait, shuffling gait, and festinating gait [19]. As typical structural damage, traumatic brain injury occurs as a consequence of a sudden acceleration or deceleration within the cranium [20]. It alters the cerebral blood flow and pressure within the skull, thereby leading to irregular physical, cognitive, and emotional behaviors. Thereby, People who have a concussion or traumatic brain injury are characterized by lower gait speed, less single-leg stance duration, and greater stride width compared with healthy ones [21, 22]. Another common presentation is stroke [23] which is affected by the blocked blood vessels of the brain and damaged cerebral circulation. Stroke is associated with a reduced stride length and cadence, increased variability in timing and amplitude of steps, and a relatively preserved arm swing [7, 24].

Although clinical researchers have quantified the typical pathological gait properties for brain disease screening, it is unknown whether some inconspicuous pathological properties are neglected. A minute alteration in gait can be **quantitatively** manifested in the relationship change between gait segments. Therefore, we envision that the gait segment relationship is a principle and generic feature for screening either single disease or compound diseases.

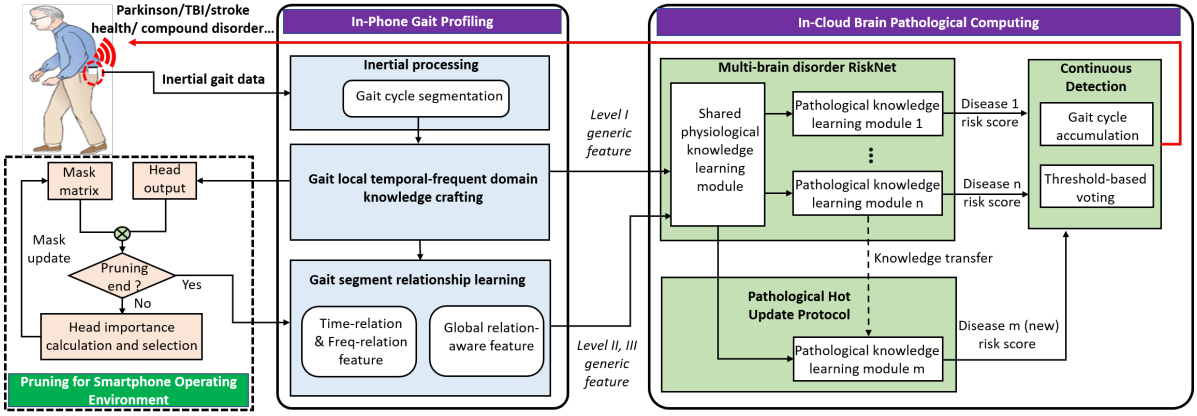


Fig. 2. The overview of *NeuralGait*, a principle and pervasive smartphone-based gait analysis system that consists of an in-phone gait profiling module and an in-cloud brain pathological computing module for multiple brain disease screening.

3 NEURALGAIT SYSTEM OVERVIEW

3.1 Design Considerations

3.1.1 Passive Sensing: Brain diseases are various and complex. Some diseases (e.g., PD) are progressive and the symptoms grow gradually. Some diseases (e.g., TBI) occur as a consequence of a sudden event. To capture the signs of these diseases in time, it is necessary to develop a passive sensing system in daily life without compromising user adherence.

3.1.2 Scalability: Brain health assessment should cover multiple types of brain diseases including neurodegeneration diseases, structural disorders, and cerebrovascular disorders, to satisfy users' various requirements. Therefore, the brain disease screening model should be general and can be scaled to newly added diseases efficiently. Therefore, the extracted features should be generic and the disease screening model should be easily updated when introducing a new disease.

3.1.3 Privacy-preserving: Privacy preservation is a critical concern for users when they use the smartphone app. As a result, we should guarantee that no sensitive information in daily activities is exposed to the third party during usage, besides the gait pathological information.

3.1.4 Energy-efficiency: As we want to provide a long-term tracking mode of brain diseases in daily life, our smartphone app computation should not occupy too many resources and can be operated with a low battery in the background process.

3.2 NeuralGait Modules

To satisfy the aforementioned considerations, *NeuralGait* is designed in a smartphone-cloud architecture, as shown in Fig. 2. The gait profiling is conducted on the smartphone to prevent sensitive information about daily activities exposed. We perform brain pathological computing and hot-update new diseases on the cloud by harnessing its powerful resources, which can enhance the energy efficiency and scalability of the system. The modules are detailed as follows:

3.2.1 In-phone gait profiling: The smartphone is put inside the user's pants pocket for capturing gait data in daily life. We first perform inertial processing to extract gait cycles and craft local temporal-frequent gait

domain features. By harnessing the importance-score-based pruning technology, we then develop and deploy the self-attention-based gait segment relationship learning model on smartphones to extract time relation features, frequency-relation features, and global relation-aware features, which are then transmitted to the cloud end.

3.2.2 In-cloud brain pathological computing: Once the gait parameters are acquired, *NeuralGait* input them to a scalable RiskNet consisting of a shared physiological knowledge module and custom pathological knowledge module to estimate the risk of each brain disease. We further design a pathological hot-update protocol based on knowledge transfer to efficiently add new diseases to the network. A continuous detection protocol is adopted to wait for gait cycle accumulation and perform voting to make a final brain health assessment.

4 IN-PHONE GAIT PROFILING

4.1 Inertial Processing

4.1.1 Smartphone Trajectory Modeling for Gait Sensing. The smartphone's built-in accelerometer captures the change rate of its linear velocity, and the gyroscope measures the angular velocity alone with one direction in the three-dimensional space. When a subject walks, he first lifts the left heel, then pushes backward on the ground to offer a counterforce as the body leans forward until the toe can be lifted in the air. The left leg swings backward so that the shank can maximize the potential energy like a pendulum. After the arrival of a point, the left leg changes to swing forward which transforms the potential energy into kinetic energy. After the left foot hits the ground, double legs are supported by the ground, and the other leg is ready to swing. These two legs alternately swing during the walk. The in-pocket smartphone's movement trajectory is like an inverted pendulum, the generated accelerations and rotations in different directions are well-captured and provide a powerful array of gait information, as shown in Fig. 3.

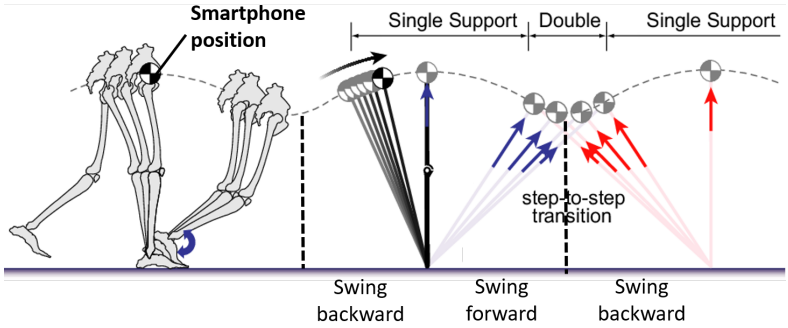


Fig. 3. Smartphone movement trajectory visualization (dashed arc) based on the inverted pendulum model [25]. The smartphone is placed in the subject's pocket around the hip (COM) while walking.

4.1.2 Gait Cycle Segmentaion. Since a smartphone may be placed in random positions inside a subject's pocket, no dimension of inertial data can fairly reflect the periodic pattern of the gait. Therefore, before extracting gait cycle for further analysis, we first calculate the magnitude signal of three-axis acceleration and gyroscope to mitigate the orientation-related noise. As the motion artifacts can cause high-frequency burrs in the raw curve, we further implement a zero-phase finite impulse response filter [10] to remove the high-frequency components. Given the walking speed is around 1.4 m/s, the bandpass filter with a passband is set from 0.75 Hz to 2.25 Hz to reduce motion artifacts. After that, we normalize the filtered magnitude signal and apply peak detection to obtain the coordinates for gait cycle segmentation.

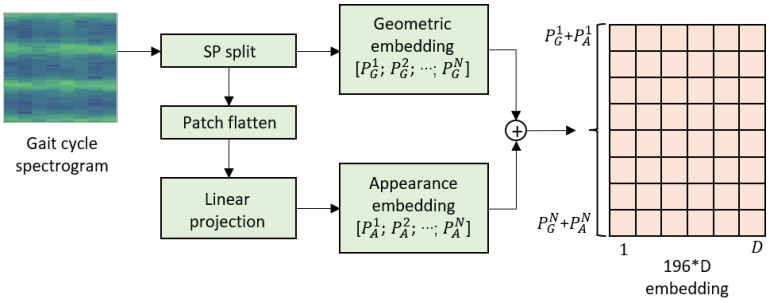


Fig. 4. A gait embedding approach based on local temporal-frequent domain knowledge.

4.2 Gait Local Temporal-Frequent Domain Knowledge Crafting

To study the local temporal-frequent knowledge of the gait, we first transform the time-series gait cycles into spectrograms. It has two advantages: 1) Augment features in both time and frequency domains; 2) low computational complexity that can be used for real-time implementation on smartphones. In detail, we segment the gait cycle signal into a series of windows with a length of 128 for each and overlapping with 64. The Hanning window function is applied to prevent the implicit pathological information from high-frequency interference and spectral leakage. After segmentation, inertial signals in these small windows are transformed into spectrograms by:

$$\mathrm{SP}\{x(t)\}(m,\omega) = |\sum_{n=-\infty}^{\infty} x[n]w[n-m]e^{-j\omega n}|^2. \tag{1}$$

Next, we evenly split the time-axis and frequency-axis of the gait cycle spectrogram into 14 segments, respectively, to generate $14 * 14$ patches. Since one gait cycle is around 0.98 to 1.07 seconds and consists of seven periods, i.e., loading response (10%), mid stance (20%), terminal stance (20%), pre swing (10%), initial swing (13%)), mid swing (14%), terminal swing (13%) [26], our divided gait patches (70 ms each) are either within a single period or at the transit boundary of two periods. Therefore, the exploration of the relationship between these gait patches can be beneficial to identify the implicit pathological properties. Then, we flatten each patch and adopt a trainable linear projection to obtain a corresponding appearance embedding with D-dimension, denoted as $\overrightarrow{P_A}$. The appearance embedding retains the local power distribution of a gait cycle. To match the local power distribution with the corresponding frequency component and time slot, the geometric embedding $\overrightarrow{P_G}$ of each patch should also be calculated. In 2D space, $\overrightarrow{P_G}$ is represented as (N_t, N_f), where $N_t \in [1, 14]$, $N_f \in [1, 14]$. To map the 2D presentation to the D-dimension presentation, we leverage the sinusoidal function for the geometric embedding, which consists of four parts, formulated as:

$$\overrightarrow{P_G} = \begin{cases} P_{G(N_t,2i)} = \sin(\frac{1}{10000^{\frac{4i}{D}}} \cdot N_t), & i \in [0, \frac{D}{4}) \\ P_{G(N_t,2i+1)} = \cos(\frac{1}{10000^{\frac{4i}{D}}} \cdot N_t), & i \in [0, \frac{D}{4}) \\ P_{G(N_f,2j)} = \sin(\frac{1}{10000^{\frac{4j}{D}}} \cdot N_t), & j \in [0, \frac{D}{4}) \\ P_{G(N_f,2j+1)} = \cos(\frac{1}{10000^{\frac{4j}{D}}} \cdot N_t), & j \in [0, \frac{D}{4}) \end{cases} \tag{2}$$

where i and j are used to distinguish between odd and even positions. After that, the geometric embedding $\overrightarrow{P_G}$ is added to the appearance embedding $\overrightarrow{P_A}$ for each patch. Finally, we obtain a $196 * D$ matrix $\mathbf{X}_P$ as **Level**

I features that contains the local temporal-frequent domain knowledge of gait cycle, expressed as: $\mathbf{X}_P = [\overrightarrow{P_A^1} + \overrightarrow{P_G^1}; \overrightarrow{P_A^2} + \overrightarrow{P_G^2}; ...; \overrightarrow{P_A^N} + \overrightarrow{P_G^N}]^T$, $N = 196$. The whole process is summarized in Fig. 4.

4.3 Gait Segment Relationship Learning

4.3.1 Motivation. Brain diseases may change the gait properties in a certain segment of a gait cycle, which will affect its relationship with other gait segments. Therefore, the relationship between different gait segments in the time and frequency domain can be an effective cue for identifying pathological patterns. The recurrent model is a typical method to learn the relationship between the consecutive data samples, but it is not able to learn the dependencies between distant positions. Consequently, we develop a learnable self-attention mechanism as gait segment relationship encoder which allows the modeling of dependencies without regard to their distance in sequences.

4.3.2 Design of Gait Segments relationship extractor. The input of the gait segments relationship extractor is matrix $\mathbf{X}_l$ that consists of an extra D-dimensional learnable "classification token" and the matrix $\mathbf{X}_P$ obtained in the previous part. This "classification token" is initialized randomly and can help learn the relations [27] in the optimization. The main component of the gait segments relationship extractor is a transformer encoder with L identical layers. Each layer consists of a multi-head self-attention sub-network and a position-wise fully connected feed-forward sub-network, with layer-norm operation applied after each sub-network [28]. The multi-head self-attention mechanism is used to jointly learn the relations between gait segments from different subspaces.

We optimize the gait segment relationship extractor by setting the learning task as predicting PD, TBI, and health subjects. As aforementioned, major neural disorders into two types can be roughly divided into two types based on the pathogenesis mechanism: neurodegeneration (e.g., PD, dementia) and structural damage (e.g., TBI, stroke, transient ischemic attack). The pathogenesis mechanism is the root of the gait relationship changes. Therefore, we hypothesize that, after optimization, this extractor gains prior knowledge of how to capture the relations between different segments within a gait cycle for neural diseases dominant by these two fundamental pathogenesis mechanisms. On the other hand, other brain diseases share similar pathogenesis mechanisms as PD and TBI, or resulted from the combination of these two fundamental pathogenesis mechanisms, so this module can serve as a principle and generic feature extractor. We do not need to further train the feature extractor even if introduces new brain diseases. As illustrated in Section 6.1, we directly feed the gait samples of stroke subjects to the frozen extractor for obtaining gait segment relationship features.

The basic idea of self-attention is to use the keys to bridge between the queries (what we are looking for) and the values (what we will actually get). Therefore, the first step is to learn the relation between the query and key for defining the attention weights to look at the data. The learnable query and key matrix are defined as: $\mathbf{Q} = \mathbf{X}_l\mathbf{W}_Q$, $\mathbf{K} = \mathbf{X}_l\mathbf{W}_K$. Since each vector-to-vector multiplication is a dot-product similarity, the relation matrix $\mathbf{R}$ can be formulated as:

$$Relation\ Matrix = \frac{(\mathbf{X}_l\mathbf{W}_Q)(\mathbf{X}_l\mathbf{W}_K)^T}{\sqrt{d_k}}, \tag{3}$$

where $\sqrt{d_k}$ is a scaling factor to make sure that the vectors won't explode. The relation matrix characterizes the relation between each pair of two patches. Based on it, we can extract the time-relation matrix and frequency-relation matrix as the **Level II features**. Time-relation matrix represents *the relationships between patches with different time gaps under each frequency component.* Frequency-relation matrix represents *the relationships between patches with different frequency gaps under each time slot.* After that, we apply the relation matrix to weight the value, followed by layer normalization, expressed as: $LN(Softmax(\mathbf{R})\mathbf{X}_l\mathbf{W}_V)$.

We employ short residual skip connections around each sub-network to allow the representations of different levels of processing to interact. With the forming of multiple paths, the network can convey the higher-level

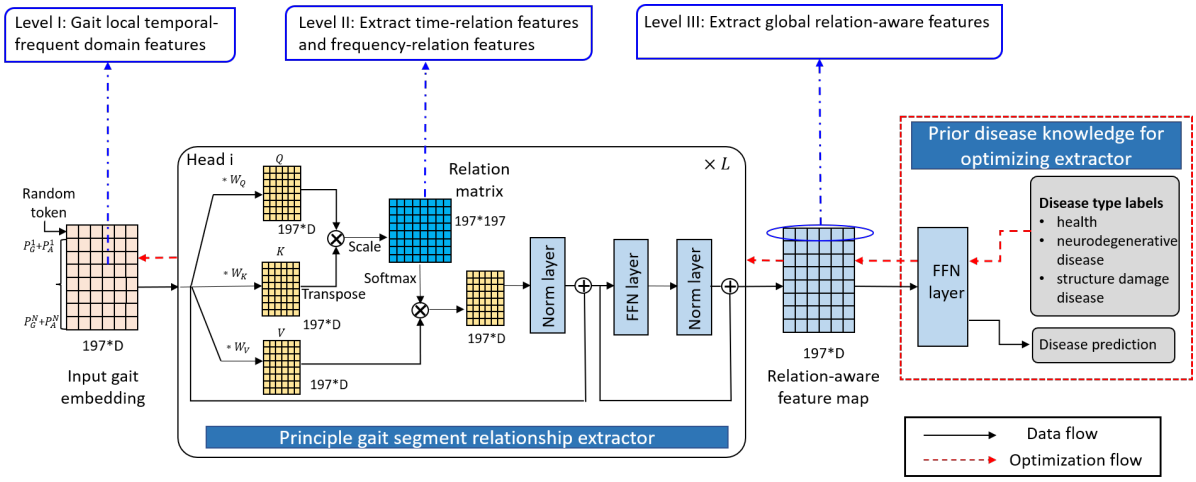


Fig. 5. A principle gait segment relationship extractor.

understanding of the last layers to the previous layers, which can re-modulate how to understand the input. After L layers learning, we can finally obtain a global relation-aware feature map, which optimizes the input embedding of the spectrogram using inner-relation knowledge. Finally, we extract the first vector (i.e., row) of the feature map as the **Level III features**, because it is a more effective feature aggregation manner that can be adaptively adjusted according to the relationship between query and key. The feature extraction process is shown in Fig. 5.

4.3.3 Gait Segment Relation Feature Visualization. We recruit 5 healthy subjects, 5 TBI subjects, 5 PD subjects, and 5 subjects with both PD and TBI for gait data collection. To mitigate the influence of demographic factors on gait data, these subjects are all male and aged 60-70. We extract 50 gait cycles from each subject. Then, a total of 250 samples for each class are fed to our optimized gait segment relationship extractor. The extracted time-relation features and frequency-relation features across different classes are visualized in this part.

To help understand the visualization result, we define some concepts. **"time gap"** denotes the number of time intervals between two patches, which varies from 0 to 12. If one patch is at time slot t_a, another patch is at time slot t_b, then the $time\ gap = |t_a - t_b| - 1$ for two patches. For example, the time gap is zero if two patches are adjacent at time slots. Similarly, **"frequency gap"** denotes the number of frequency intervals between two patches, which varies from 0 to 12. If one patch is at frequency component f_a, another patch is at time slot f_b, then the $frequency\ gap = |f_a - f_b| - 1$ for two patches. **"Self"** represents the patch itself.

Time-relation analysis: Fig. 6 shows the relationships between patches at different time gaps. As aforementioned, a gait cycle spectrogram is evenly divided into $14 * 14$ patches, and each patch can be denoted as (t, f) in spatial. Each element in the time-relation heatmap represents the relation value between patch (t_a, f) and patch (t_b, f), where $|t_a - t_b| - 1$ corresponds to time gap value in x-axis, and f corresponds to frequency component value in y-axis. We can observe that gait patches at adjacent time slots are generally related more than those at distant time slots. This observation is dominated by the gait control theory, so the time-relation heatmaps among the four classes share almost the same pattern, which makes the pathological-caused time-relation differences subtle. We leverage red ellipses to highlight certain differences among these four classes in Fig. 6. Table 1 further lists all these highlights and provides quantitative relation values for better understanding. For example, at frequency component 1, the relation between time-adjacent patches is distinct between TBI subjects and other subjects.

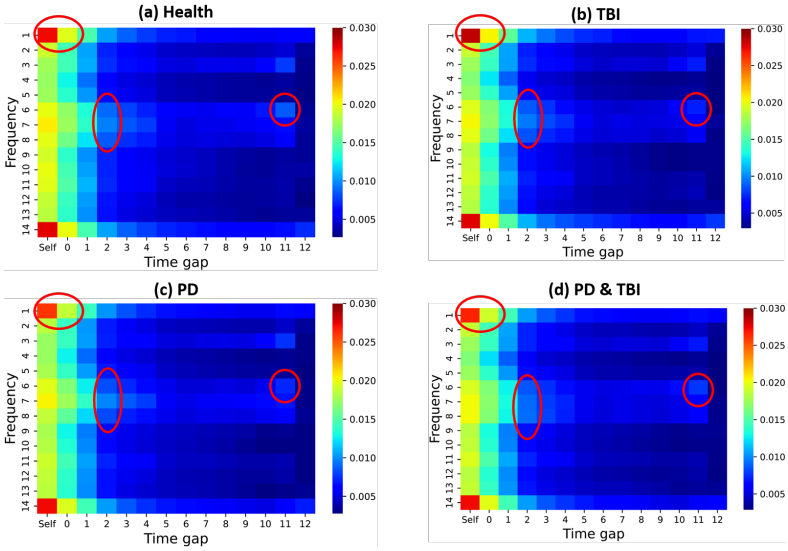
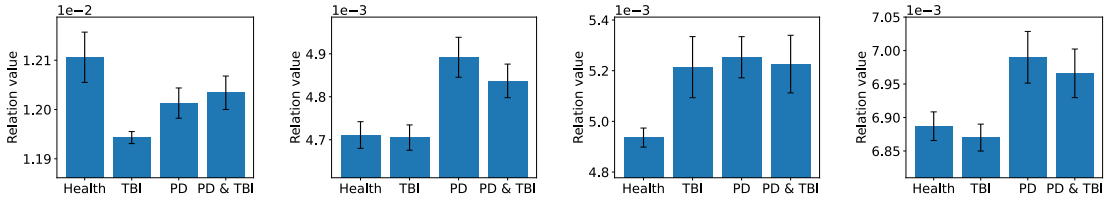


Fig. 6. The comparison of time-relation heatmaps among four classes. Each element in the heatmap represents the relation value between two patches which are both at frequency y but gapping x time slots in the original spectrogram. The red ellipses highlight the difference between these four classes.

At frequency components 6, 7, 8, the relation between patches with relatively small time gaps can be served as a cue to identify a single neural disease and compound neural diseases. For further exploration, we average the time-relation values across the frequency axis, and then study the difference in relation values among four classes, as shown in Fig. 7. We have the following findings:

- Health subjects show higher relations between gait patches at adjacent time slots (i.e., time gap= $0 \sim 1$) than subjects with brain diseases. It indicates that healthy subjects have higher walking coherence within 150ms, and seldom encounter abrupt gait changes.
- PD subjects show higher relations between gait patches at distant time slots (i.e., time gap= $2 \sim 11$) than healthy and TBI subjects. That is to say, for PD patients, different gait phases (such as the stance phase and swing phase) show higher dependency on each other.
- TBI subjects show the lowest relations between gait patches at different time gaps overall, but show high relations between patches at the start time slot and end time point. We obtain that TBI patients are worse in gait coherence and stability, even though the start and end stages of a gait cycle may be smooth.

Frequency-relation analysis: We demonstrate the relationships between patches with different frequency gaps in Fig. 8. It is observed that gait patches with small frequency gaps are generally related more than those with large frequency gaps. Such a trend dominates the distribution of the frequency-relation heatmaps, so the pathological-caused frequency-relation differences are inconspicuous visually among the different four classes. We use red ellipses to highlight slight differences in frequency-relation heatmap among four types of subjects. Table 1 lists these highlighted differences on heatmaps and present quantitative relation results among four classes. For example, at time slot 2,3,4, the relation between patches with relatively small frequency gaps is different for TBI and PD. For further exploration, we average the frequency-relation values across the time axis, and then study the difference in relation values among four classes, as shown in Fig. 9. We have the following findings:



(a) Time gap = 0 ∼ 1 (b) Time gap = 2 ∼ 11 (c) Time gap = 12 (d) Overall time gap

Fig. 7. Relations between patches at different time gaps among four classes.

Table 1. The summary of ellipse highlights in relation heatmaps among four classes.

Ellipse Highlight	Patch 1 Coordinate	Patch 2 Coordinate	Relation value				Cue
			Health	TBI	PD	PD&TBI	
(0,1)	(t, 1)	(t+1, 1)	1.93e−2	**2.12e−2**	1.86e−2	1.89e−2	Highest: TBI
(2,8)	(t, 8)	(t+3, 8)	8.46e−3	8.49e−3	8.16e−3	**9.15e−2**	Highest: PD & TBI
(11,6)	(t, 6)	(t+12, 6)	**8.44e−3**	7.07e−3	7.19e−3	7.48e−3	Highest: Health
(8, self)	(8, f)	(8, f)	**2.26e−2**	**2.25e−2**	2.05e−2	2.01e−2	close: Health and TBI
(3, 1)	(3, f)	(3, f+2)	**9.16e−2**	**9.11e−2**	8.57e−2	8.35e−2	close: Health and TBI

- Overall, TBI subjects show the highest relations between gait patches with different frequency gaps, and PD subjects show the lowest relations between gait patches with different frequency gaps.
- Healthy subjects show lower relations (near the same as PD) between gait patches with 7 ∼ 9 frequency gaps, compared with other numbers of frequency gaps.

4.4 Pruning for Smartphone Operating Environments

As aforementioned, learning the relationship between gait segments needs performing scaled dot-product attention multiple times in parallel, which can introduce much memory consumption and computational cost. As a result, it is hard to deploy such a gait segment relationship extractor on off-the-shelf smartphones. To solve this issue, our idea is to remove some unimportant attention heads in the self-attention model that has been trained without compromising the prediction performance.

We first construct a 12×12 mask matrix to mask all the 144 attention heads in 12 layers, where each mask variable δ_h belongs to $\{0, 1\}$. Then, we multiply the outputs of trained 144 attention heads by the mask matrix. At the beginning of the pruning process, all mask variables are set as 1. The expected sensitivity of the encoding model to the mask variables is formulated as [29]:

$$S_h = \mathbb{E}_{x \sim X} \left| \frac{\partial \mathcal{L}(x)}{\partial \delta_h} \right|, \tag{4}$$

where $\mathcal{L}(x)$ is the loss on input inertial data sample, X is the data distribution. Higher S_h indicates that changing δ_h is prone to have large effect on the model. Therefore, S_h can be regarded as the importance score of each attention head. After obtaining the importance score of each head, we select the top 5 least important heads, and set the corresponding mask variables as 0. As a result, the self-attention model is updated. We iteratively perform the pruning process to mask unimportant heads and update the self-attention model, until the prediction

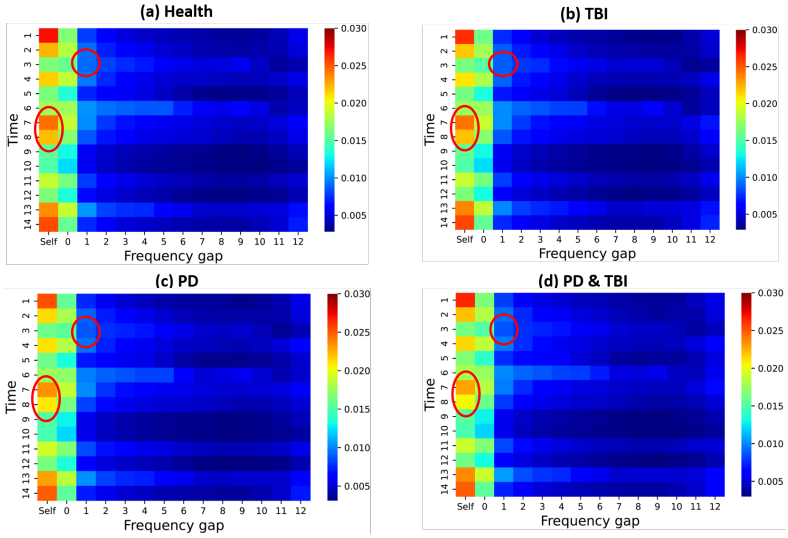
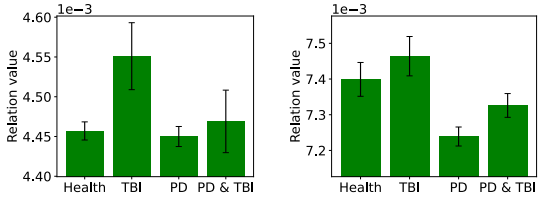


Fig. 8. The comparison of frequency-relation heatmaps among four classes. Each element in the heatmap represents the relation value between two patches which are both at time y but gapping x frequency components in the original spectrogram. The red ellipses highlight the difference between these four classes.



(a) Frequency gap = 7 ∼ 9 (b) Overall frequency gap

Fig. 9. Relations between patches with different frequency gaps among four classes.

performance starts to drop. Finally, we obtain a lightweight gait segment relationship extractor model that can be deployed on smartphones.

5 IN-COULD BRAIN PATHOLOGICAL COMPUTING

5.1 Multi-Brain Disorder RiskNet

After achieving the Level I, II, III gait features from the smartphone end, we develop a pathological risk estimation network (RiskNet) to detect multi-label brain diseases, as shown in Fig. 10. The design of RiskNet should satisfy two requirements: 1) one subject possibly has more than neural diseases, these diseases are not mutually exclusive, so our model needs to detect multi-label diseases simultaneously; 2) Our model needs a scalable architecture for adapting to the new involved disease with minimum efforts. Therefore, our developed RiskNet consists of a shared physiological knowledge module and multiple custom pathological knowledge modules (branch), where each branch is used for predicting one type of neural disease, these prediction branches are not mutually exclusive.

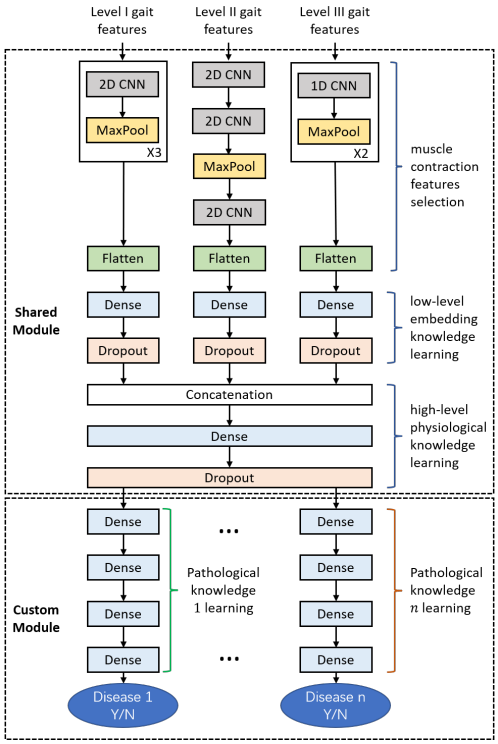


Fig. 10. A scalable RiskNet architecture for multiple brain diseases assessment.

5.1.1 Electrophysiological Neural Basis. Human gait depends on a complex interplay of the nervous system and musculoskeletal system. Information in the brain is disseminated through neuron communication, which is based on action potentials and chemical neurotransmitters. The action potential is controlled by the concentration of intracellular and extracellular ions [30], formulated as:

$$E_m = \frac{RT}{F}\left(\frac{\sum_i^N P_{M_i^+}[M_i^+]_{out} + \sum_j^M P_{A_j^-}[A_j^+]_{in}}{\sum_i^N P_{M_i^+}[M_i^+]_{in} + \sum_j^M P_{A_j^-}[A_j^+]_{out}}\right), \tag{5}$$

where E_m is the membrane action potential, P is the selectivity for that ion, R is the gas constant, T is the absolute temperature, F is Faraday's constant. When an action potential reaches the presynaptic terminal, neurotransmitters are released from the neuron into the synaptic cleft [31]. These neurotransmitters then bind to receptors on the postsynaptic side and open particular types of ion channels that span the membrane, which generates action potential again [31]. In this way, neural excitation is propagated through the whole body. As the motor neurons stimulate the muscle fibers, the generated action potential influences intracellular Ca^{2+} concentration. The Ca^{2+} ions flood into the muscle cell to bind with troponin, and contract using ATP as energy [32], which finally produces force and initiates walking.

5.1.2 Shared Physiological Knowledge Module. Understanding the neural controlling gait theory is fundamental for investigating the association between neural pathology and gait impairment, but it is hard to model the aforementioned process to quantify neural features. Therefore, we design a data-driven module to learn the fundamental physiological-level neural knowledge from the gait domain features and relation features, which

are shared by each pathological learning module. We first fine-tune each level of gait features to extract muscle contraction information. In detail, Level I gait domain features ($197 * 768 * 1$) pass through three blocks consisting of a 2D convolutional layer and a pooling layer; Level II time-relation and frequency-relation features ($14 * 14 * 2$) go though two 2D convolutional layers, a pooling layer, and a 2D convolutional layer, respectively; Level III global relation-aware features ($768 * 1$) pass through two blocks consists of a 1D convolutional layer and a pooling layer. After that, a dense layer and a dropout layer are applied to each type of fine-tuned feature to learn the low-level embedding knowledge. Finally, all this low-level knowledge is concatenated together and goes through another dense layer and dropout layer to extract high-level physiological knowledge, such as neurons excitation and propagation.

5.1.3 Custom Pathological Knowledge Module. Since different brain disorders have their typical pathological characteristics, we design custom modules to learn each type of brain pathological knowledge from shared physiological knowledge. In each custom module, we adopt four dense layers and a softmax activation layer. The softmax layer is applied for twofold reasons. First, the softmax function enhances the discrepancy of predicted probability scores between different classes of input. Second, the differentiability of the softmax function can enable the backpropagation to update weights. The cross-entropy is used as a loss function to penalize the deviation of the model's output pathological probabilities from desired probabilities. Finally, for each gait cycle, we can obtain the predicted risk scores and pathological presentation (i.e., yes or no) on every disease stored in the cloud.

5.2 Pathological Hot Update Protocol

Brain diseases are complex and various. Clinical achievements towards brain diseases are always updated along with time. How to add the new disease in the RiskNet for onset risk estimation with minimal effort is a challenge. In addition, we should consider that the labeled data of the patients in real life is insufficient, specifically for some under-researched brain diseases. Therefore, we develop a pathological hot update protocol using transfer learning technology.

In detail, the gait segment relationship extractor and physiological knowledge learning module are frozen, while leaving the pathological knowledge learning module open for optimization. According to the characteristics of the newly added disease, a current pathological learning module designed for the same disease category is selected as the source domain, which is then trained with new disease data to better adapt to the target domain. Worth mentioning, even if the target domain has a small amount of patient data or the similarity between the source and target domains is not large, the dense layer in the pathological learning can guarantee the transferring effectiveness due to its capacity to preserve model complexity [33]. In this process, similar pathological knowledge is transferred to the new disease module and largely reduces learning efforts. Finally, such a new pathological knowledge learning module is saved in the cloud end and the RiskNet is updated.

5.3 Continuous Detection Protocol

Brain disease has multiple types. Progressive neurodegenerative disease is a common type of brain disease. For these patients in the early stage, symptoms are barely noticeable and the pathological gait representation may only occur randomly during walking. On the other hand, for healthy people, some momentary random events (e.g., stumbles, twisting, spastic) can not be avoided in daily life, which will cause false alarms and compromise prediction performance. Therefore, gait data accumulation is better to assess the user's condition and make a more reliable prediction. Smartphone-based gait sensing modality provides continuous and passive tracking to achieve this goal with no burden in daily life.

As aforementioned, we have predicted the disease presentation for each gait cycle. Given a sequence of gait cycles (samples) $\{S_1, S_2, ..., S_K\}$ from one user, by leveraging an adaptive threshold-based voting method, the final

decision on whether this person has brain disease j is formulated as:

$$P_j = \begin{cases} \text{Yes} & \frac{1}{K}\sum_{k=1}^{K} p_j^k > \gamma \\ \text{No} & \text{Otherwise} \end{cases}, \tag{6}$$

where p_j^k=1 if the gait sample S_k is predicted to have pathological presentation of disease j, otherwise p_j^k=0; γ is an adaptive threshold that we will discuss the optimal selection in Section 6.3. For the cumulative gait cycles, we provide a final prediction result, i.e., yes or no, on each brain disease stored in the cloud. Finally, the users can get a more reliable assessment on their brain health.

6 EVALUATION

6.1 Experiment Setup

6.1.1 Participants Recruitment. Our study is approved by the Western Institutional Review Board (WIRB). In a 3-month study, we online recruit 988 healthy people and 417 patients (with a single or combination of PD, TBI, or stroke) from the U.S. to join our research. We cooperate with the professional medical institute to conduct clinical diagnoses for these subjects to confirm their health condition and exclude subjects with physical issues in walking. Among them, 66% are males and 34% females with their ages ranging from 20 to 78. The average age is 61.

6.1.2 Data Collection. We collect the subjects' data in an uncontrolled daily-life environment. Three smartphone models are involved in the data collection, which are iPhone 6 plus, iPhone 6, and iPhone 5s. Subjects need to install our smartphone App for gait data collection. Specifically, we instruct each subject to put the smartphone inside the left-back pocket of the pants and walk straight for at least 20 steps with no stops to complete each recording at the home. Note that all patients are requested to identify their medication intake situations for each walking recording. To reduce the interference caused by the treatment, we only select the walking recordings "just before any medication" and "never have medication". No other sensor data are discarded apart from the medication reason. Since our system is performed based on the gait cycle, each walking recording is then segmented into multiple gait cycles. Finally, we obtain 56089 healthy samples (gait cycles), 23207 PD samples (gait cycles), 11920 TBI samples (gait cycles), and 6051 stroke samples (gait cycles).

6.1.3 Partition. Our system is designed for unseen users in the real world, so we employ a person-independent training/test data split protocol. To examine our system's feasibility to screen multiple neural diseases, the data partition (summarized in Table 2) has two parts: ***1) Train a generic model:*** We use the gait cycles of the PD, TBI, and healthy subjects to train a generic model, which is divided into two independent subsets. One subset (around 35000 samples) is used for training the gait segment relationship extractor. After that, the gait data from the other subset are input to the optimized extractor for obtaining the gait features, which are then used for training and testing the RiskNet. Among them, two-thirds of the subjects (around 24000 samples) are used for training and the remaining subjects (around 12000 samples) are used for testing. ***2) Train a customized model:*** When introducing a new disease, the gait segment relationship extractor and the shared module of RiskNet are frozen, while only leaving the customized module (i.e., pathological knowledge learning module) open for optimization. As stroke is associated with structural damage and shares certain pathological characteristics with TBI, we train the customized pathological knowledge learning module for stroke detection by leveraging transferring learning technology. The stroke samples are divided into two person-independent subsets. Around 4000 stroke samples and 12000 non-stroke samples (randomly selected from the training set of the generic model) are used for training the stroke prediction branch. The remaining 2000 stroke samples are used for testing. In total, we have around 14000 gait samples with a single or combination of PD, TBI, or stroke to test our system performance.

6.1.4 Implementation and Training policy. We implemented our models in PyTorch. To reduce the influence of the unbalanced data, weighted losses are used for each branch and across the branches. The attention-based gait segment relationship extractor is trained with the AdamW optimizer. After that, the extractor is optimized and converted to PyTorch mobile model for deployment on the phone. To train the RiskNet, a stochastic gradient descent optimizer with momentum is adopted with a 2e-5 learning rate, 0.9 momentum, and 0.01 weight decay. The learning rate decays over time with a multiplicative factor of 0.9 and a step size of 10.

Table 2. Data partition list.

	Training set	Testing set
Generic model	35000 samples for extractor, 24000 samples for RiskNet	12000 samples
Customized model	16000 samples	2000 samples

Table 3. *NeuralGait*'s multi-branch output and corresponding interpretation.

Output	Interpretation
[0, 0, 0]	Health
[1, 0, 0]	Only PD
[0, 1, 0]	Only TBI
[0, 0, 1]	Only Stroke
[1, 1, 0]	PD & TBI
[1, 0, 1]	PD & Stroke
[0, 1, 1]	TBI & Stroke
[1, 1, 1]	PD & TBI & Stroke

6.2 Evaluation Metrics

In the brain health assessment application, healthy subjects and diseased subjects are mutually exclusive, but one subject possibly suffers from more than one type of neural disease, which makes the disease detection branches not mutually exclusive. Therefore, our evaluation is specifically designed for the multi-label multi-class disease classification system.

With sample X_i as input, we define the ground truth as $Y_i = [Y_i^{\lambda_1}, ..., Y_i^{\lambda_k}]$, and prediction as $Z_i = [Z_i^{\lambda_1}, ..., Z_i^{\lambda_k}]$, where

$$Y_i^{\lambda} = \begin{cases} 1 & \text{If the actual output of disease } \lambda \text{ branch is Yes} \\ 0 & \text{Otherwise} \end{cases}, \tag{7}$$

$$Z_i^{\lambda} = \begin{cases} 1 & \text{If the predicted output of disease } \lambda \text{ branch is Yes} \\ 0 & \text{Otherwise} \end{cases}, \tag{8}$$

$\lambda_1, ..., \lambda_k$ represent multiple neural disease types that our system can detect. In our evaluation, we examine the screening performance for three diseases, i.e., PD, TBI, and stroke. Table 3 shows possible outputs of our system. A set of metrics are selected to evaluate *NeuralGait*'s capability as follows:

6.2.1 Overall Multi-Label Multi-Class Classification. We use exact matching accuracy and global accuracy to examine *NeuralGait*'s overall performance in screening multiple neural diseases simultaneously.
Exact Matching Accuracy: Exact Matching Accuracy (EMA) ignores partially correct predictions (consider them as incorrect). It measures the *NeuralGait*'s capability to identify all neural diseases with no misses, formulated as:

$$Exact\ Matching\ Accuracy = \frac{1}{n}\sum_{i=1}^{n} P(Y_i = Z_i), \tag{9}$$

where P is the indicator function [34], exactly matching will return 1.
Global Accuracy: Global accuracy distinguishes partially correct predictions and complete correct predictions. The accuracy for each sample is defined as the proportion of the predicted correct disease labels to the total number (predicted and actual) of disease labels [35]. Then, the global accuracy is the average across all samples, formulated as:

$$Global\ Accuracy = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Y_i \cap Z_i}{Y_i \cup Z_i}\right|. \tag{10}$$

6.2.2 Healthy VS Unhealthy. We evaluate the system's performance in detecting healthy subjects (normal) and subjects with neural diseases (abnormal). The positive class is defined as healthy subjects (i.e., $Y_i = [0, 0, 0]$), and the negative class is defined as subjects with neural diseases. These two classes are mutually exclusive, so we use recall, precision, and F1-score as binary classification metrics.

6.2.3 Specific Neural Disease Type Screening. For each disease prediction branch λ, the positive class is defined as the samples with disease λ, and the negative class is defined as the samples without disease λ that consist of healthy samples and other diseased samples. Note that if a sample contains but is not limited to disease λ, it is still considered as the positive class of disease prediction branch λ. We define the sensitivity$^\lambda$, FAR$^\lambda$, and F1-score$^\lambda$ for screening each type of neural disease [34].
Sensitivity$^\lambda$: Sensitivity$^\lambda$ measures whether the *NeuralGait* is sufficiently sensitive to pick up a neural disease λ with no misses. It is the fraction of true positive numbers to the actual number of disease λ, which refers to the probability of a positive assessment result for patients with a disease λ, given by:

$$Sensitivity^\lambda = \frac{\sum_{i=1}^{n} Z_i^\lambda Y_i^\lambda}{\sum_{i=1}^{n} Y_i^\lambda} \tag{11}$$

FAR$^\lambda$: False alarm rate (FAR$^\lambda$) measures the *NeuralGait*'s ability in avoiding false diagnoses of disease λ among healthy subjects. It is the fraction of the false positive number to the actual number of subjects without disease λ, formulated as:

$$FAR^\lambda = \frac{\sum_{i=1}^{n} Z_i^\lambda - \sum_{i=1}^{n} Z_i^\lambda Y_i^\lambda}{n - \sum_{i=1}^{n} Y_i^\lambda} \tag{12}$$

F1-score$^\lambda$: F1-score$^\lambda$ is the harmonic mean of sensitivity$^\lambda$ and precision$^\lambda$. It is non-sensitive to the distribution of subjects with and without disease λ and can be formulated as:

$$F1\text{-}score^\lambda = \frac{2\sum_{i=1}^{n} Z_i^\lambda Y_i^\lambda}{\sum_{i=1}^{n} Y_i^\lambda + \sum_{i=1}^{n} Z_i^\lambda} \tag{13}$$

6.3 Disease Screening Performance

6.3.1 Threshold Searching. To maximize the applicability of *NeuralGait* in real-world scenarios, it is important to study how many steps a user needs to walk for achieving a more reliable brain health assessment result. To determine the required cycle counts and the voting threshold for each time assessment, we leverage the

macro F1-score$^{\lambda}$ to search for the optimal solution, where 1/3 of users' data is used for searching, and other 2/3 users' data is used for testing. Note that macro F1-score$^{\lambda}$ is computed using the arithmetic mean of per-class (i.e., disease$^{\lambda}$ class and non-disease$^{\lambda}$ class) F1-score, which is different from the F1-score$^{\lambda}$. We only use *macro* F1-score$^{\lambda}$ for threshold selection but not for testing the screening performance. After searching, we select 14 as our unified gait cycle counts for each time brain health assessment. The voting thresholds for screening PD, TBI, and stroke are set as 5/14, 3/14, and 4/14, respectively.

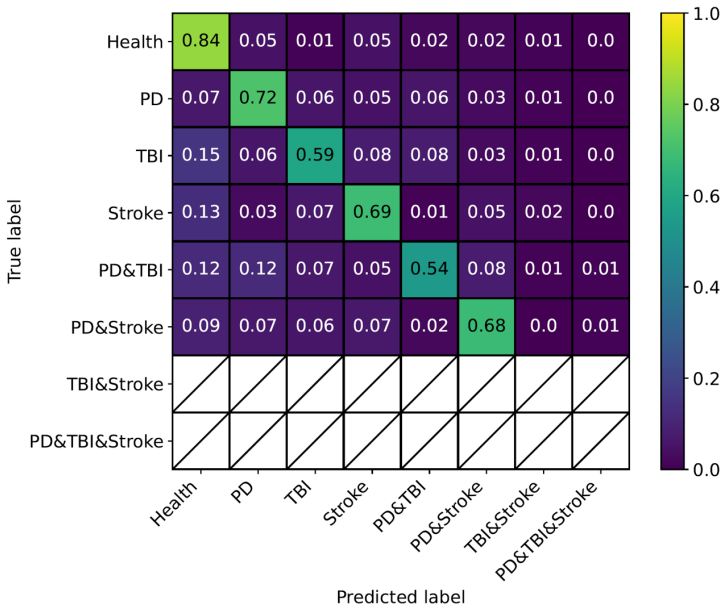


Fig. 11. Normalized confusion matrix for multi-label multi-class neural disease classification.

6.3.2 Results. With the current voting threshold setting, Table 4 presents all the evaluation results. We analyze the *NeuralGait*'s performance mainly in three aspects.

Overall Multi-Label Multi-Class Classification: In general, *NeuralGait* achieves 77.7% exact matching accuracy and 80.1% global accuracy for the multi-label multi-class neural diseases classification. We further demonstrate the normal confusion matrix of neural disease classification to provide a detailed view. The x-axis is the predicted multi-branch output, and the y-axis is the actual multi-branch output. Note that none of the recruited subjects have "TBI & Stroke" or "PD & TBI & Stroke", so the last two lines in our confusion matrix are null. As shown in Fig. 11, we have the following observations:

- 16% of the healthy samples are predicted as diseased samples, where PD predictions take up to 9%.
- For subjects who have a single neural disease, 59%-72% of the screening results are completely correct, and 8%-9% of the screening results contain but are not limited to the ground truth disease.
- For subjects who have multiple neural diseases simultaneously, 54%-68% of the screening results are completely correct, and more than 17% of the screening results are partially correct.

Health VS Unhealthy Detection: *NeuralGait* obtains 88.9% F1-score to classify healthy subjects and subjects with neural diseases. The recall and precision reach up to 84.1% and 93.4%.

Specific Neural Disease Type Screening: *NeuralGait* reaches more than 70% F1-score of screening each type of neural disease on unseen subjects. More specifically, *NeuralGait* is most sensitive to picking up PD and achieves the lowest false alarm rate when screening TBI. This is consistent with our design consideration. As a neurodegenerative disease, PD symptoms grow gradually over time, so a more sensitive screening approach can help identify the risk early for in-time treatment. On the other side, TBI occurs as a consequence of a sudden event and the symptoms are noticeable and more serious at first, so the false alarm rate becomes *NeuralGait*'s highest priority to improve user experience, instead of sensitivity. Worth mentioning, even though the stroke prediction branch is trained from the customized model, instead of training from the very beginning, it still can provide reliable stroke screening performance with 71.4% sensitivity and 8.9% false alarm rate. To conclude, *NeuralGait* can be used for reliable brain health assessment on unseen users in the real world.

Table 4. *NeuralGait*'s performance in a) distinguishing healthy subjects and diseased subjects; b) screening specific neural disease types; c) overall multi-label multi-class detection.

	Health		PD	TBI	Stroke
F1-score	88.9%	F1-score$^{\lambda}$	76.8%	70.2%	71.4%
Recall	84.1%	Sensitivity$^{\lambda}$	79.1%	65.5%	76.1%
Precision	93.4%	FAR$^{\lambda}$	9.7%	5.9%	8.9%
Exact matching accuracy			77.7%		
Global accuracy			80.1%		

6.4 Disease Progression Analysis

Neurodegenerative disease (e.g., PD) is a common type of brain disease and its symptoms are rarely noticeable at first but grow gradually over time. A recent study shows that arrhythmic gait pattern and freezing of gait occur frequently with the progression [36]. Early detection of neurodegenerative disease can help in-time treatment to slow or halt disease progression. Therefore, we are motivated to investigate whether *NeuralGait* can be used to effectively screen the PD in the early stage.

Based on the previous clinical study[37], we classify the PD as early stage and middle stage according to the onset years, i.e., early stage refers to the patients who have been onset within three years, and middle stage refers to the patients who have been onset more than three years. Fig. 12 shows the predicted PD risks of patients under different onset years. Note that, the predicted risk of each subject is calculated by the average predicted disease label values (i.e., 0 or 1) among his/her contributed samples. We observe that the average predicted risks rapidly increase to above 0.85 as the onset years reach up to three years. Compared with the middle stage, the range of prediction risk is slightly larger in the early stage, but the average predicted risks does not display much disparity.

In addition, we examine the F1 score and sensitivity for screening PD in the early stage and middle stage. The performance between the patients with only PD and with at least PD is further compared to validate system's resilience to compound diseases. Fig. 13a shows that the F1-score in early-stage screening is about 5% higher than that in middle stage for both types of PD patients. As observed in Fig. 13b, the screening sensitivity is above 82% for both types of PD patients in the early stage, whereas the sensitivity drops to around 71% in the middle stage. These results strongly indicate that *NeuralGait* is more sensitive and robust to PD patients in the early stage, which can be used for reliable early screening in daily life.

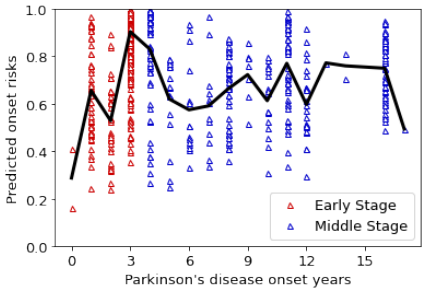
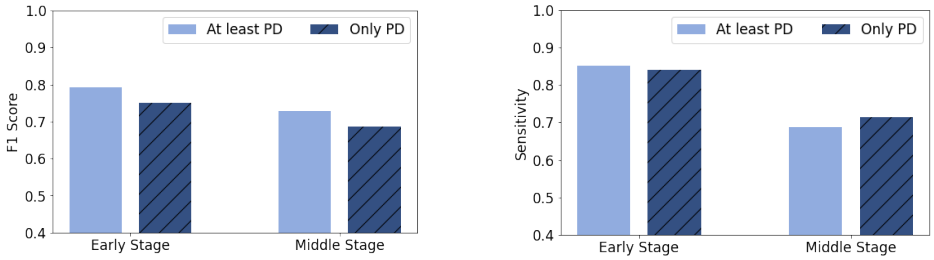


Fig. 12. The prediction of PD onset risks for patients under different onset years.



(a) (b)

Fig. 13. The F1 score and sensitivity for screening PD in early stage and middle stage.

6.5 Smartphone Overhead Analysis

We evaluate the system overhead on the smartphone end, which contains the feature extraction functions based on attention-based encoder. The evaluations are performed on four smartphones (released year = 2017-2020) with different hardware. The models are implemented on Pytorch mobile.

The averaged runtime of a single gait cycle, averaged CPU load, and battery usage are measured. Each phone continuously ran with 600 gait cycle inputs to have reliable measures. The results are reported in Table 5. The run time of a cycle ranges from 726 ms to 837 ms, with a reasonably low average CPU usage. Consider that each gait cycle takes about one second on average. The system is capable of processing data in real time. Among the four devices, the Pixel 2 consumes the most battery, even in which case, a 14-gait-cycle screening will only use 13.3 mAh (0.5% of its battery capacity). These results indicate the battery use of our model is well controlled via pruning.

Table 5. System overhead on different smartphones.

Phone Model	Runtime (ms)	Avg CPU (%)	Battery Usage (mAh)
Galaxy Note20	726	26	465
Galaxy S20+	837	29	128
Pixel 2	687	20	570
Pixel 4	631	18	249

7 INCLUSIVENESS STUDY

7.1 Impact of Gender

Females and males have different gait patterns due to their different musculoskeletal structures. Females often walk with a higher cadence and a shorter stride length compared with males. In addition, women's gait is characterized by a greater ankle range of motion while men tend to have a greater hip range of motion [38]. Therefore, it is important to examine whether the gender-induced gait difference will affect *NeuralGait* performance. Fig. 14 shows that the screening sensitivity of PD and stroke is nearly the same for both males and females while displaying a slight difference in screening TBI, i.e., 64.5% sensitivity among males and 67.2% sensitivity among females. For each type of disease screening, females are associated with a slightly higher false alarm rate compared with males. Overall, *NeuralGait* can achieve above 64% sensitivity and less than 12% false alarm rate on both males and females, which validates its inclusiveness to gender. Furthermore, our results indicate that gender is an independent biological factor that can help personalize brain health assessment.

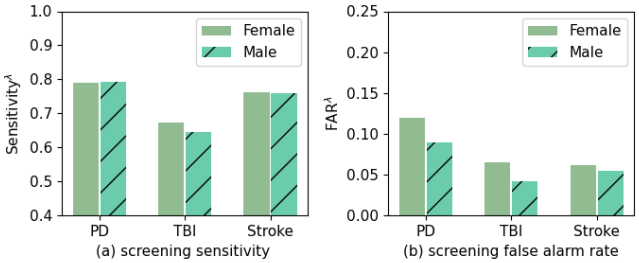


Fig. 14. The impact of gender on *NeuralGait*'s performance of screening each type of neural disease.

7.2 Impact of Age

Age is an inevitable factor that will affect gait patterns. Aging people shows lower walking speed and decreased hip range of motion due to the decline in muscle strength [39]. Consequently, we investigate how the age factor affects *NeuralGait* performance in screening diseases. Based on two age thresholds that well capture the impact of age on walking speed, subjects are divided into three age groups, i.e., <50, 50-70, >70. The results are shown in Fig. 15. Due to the lack of training data in the <50 age group, *NeuralGait* is not sensitive to pick up neural diseases for younger subjects. Among the 50-70 age group, we observe that *NeuralGait* achieves above 71.4% sensitivity in identifying each type of neural disease. For the subjects aging above 70, the TBI screening sensitivity and stroke screening sensitivity drops more than 12%, whereas the PD screening sensitivity decreases slightly. The false alarm rate of screening each type of neural disease does not show a notable ascending or descending trend across different age groups, besides screening stroke among subjects aging below 50. These results indicate that age factor is possible to influence system performance, particularly for stroke screening. This is because stroke prediction is developed based on a customized model adapted from other disease prediction branches, where the discrepancy between age-induced and stroke-induced gait properties is not well learned. In the future, we will optimize the customized model design to improve the age inclusiveness of *NeuralGait*.

7.3 Impact of Smoking

Long-term smoking has a negative effect on the nervous system and brain. The compound in tobacco can provoke white blood cells in the central nervous system to attack healthy cells, which will result in neurological impairment.

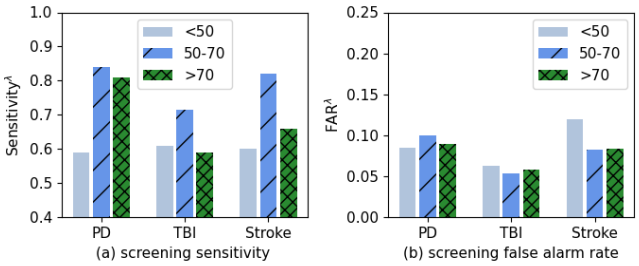


Fig. 15. The impact of age on *NeuralGait*'s performance of screening each type of neural disease.

Therefore, we are motivated to investigate whether this impairment will influence the screening performance of brain diseases. As observed in Fig. 16, *NeuralGait* achieves 8%-10% higher disease screening sensitivity among non-smokers than those among smokers. The false alarm rate for TBI screening and stroke screening is almost the same among smokers and non-smokers, while the false alarm rate for PD screening displays a slight increase among smokers. These results indicate that smoking can affect the *NeuralGait*'s sensitivity in brain health assessment, which is consistent with the smoke-induced neural damage theory. In the future, we plan to consider the smoking factor in *NeuralGait*'s development.

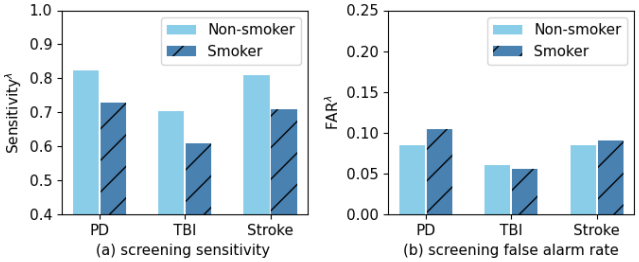


Fig. 16. The impact of smoking on *NeuralGait*'s performance of screening each type of neural disease.

7.4 Impact of Smartphone Type

NeuralGait leverages the smartphone built-in accelerometer and gyroscope to collect gait. For different types of smartphones, the sensor type, sensor deployment, and sensor quality are various. As a result, we are curious about if *NeuralGait* is sensitive to smartphone models. In the training and testing phase, the gait data are collected from three phone models, i.e., iPhone 6 plus, iPhone 6, and iPhone 5s. As shown in Fig. 17, although the system performance varies across different smartphone types, *NeuralGait* can provide a reliable assessment with sensitivity above 75.3% for PD, above 62.5% for TBI, and above 74.2% for stroke. The false alarm rate is lower than 12% for identifying each type of neural disease. To conclude, *NeuralGait* is resilient to different smartphone models.

8 GENERALIZATION STUDY

To evaluate the generalization performance of our proposed gait analysis approach for neural disease screening, we perform an effectiveness analysis and compute efficiency analysis on new unseen disease screening.

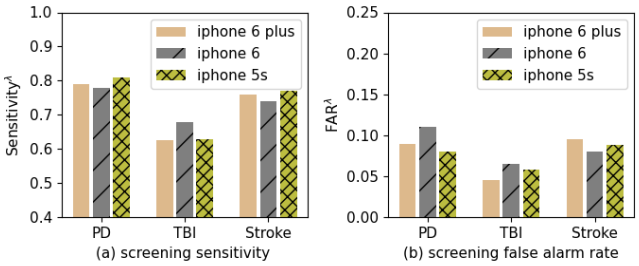


Fig. 17. The impact of smartphone models on *NeuralGait*'s performance of screening each type of neural disease.

8.1 Effectiveness Analysis

To examine the effectiveness of our developed generic gait feature extraction approach, we perform the feature correlation analysis between the subjects with the same disease, and between the subjects with different diseases, respectively.

Specifically, we recruit 10 stroke patients, 5 Alzheimer's patients, and 5 healthy subjects for gait data collection. To reduce the gender and age impacts on gait data, these subjects are all male and aging in the range of 60-70. We extract 10 gait cycles (samples) from each subject. Then, each sample is directly fed to our developed gait segment relationship extractor and the shared physiological knowledge learning module successively (where the parameters are frozen). The obtained global relation-aware features and shared physiological features are used in our evaluation. The feature correlation analysis is performed on three groups: 1) 5 stroke subjects VS another 5 stroke subjects; 2) 5 stroke subjects VS 5 Alzheimer subjects; 3) 5 stroke subjects VS 5 healthy subjects.

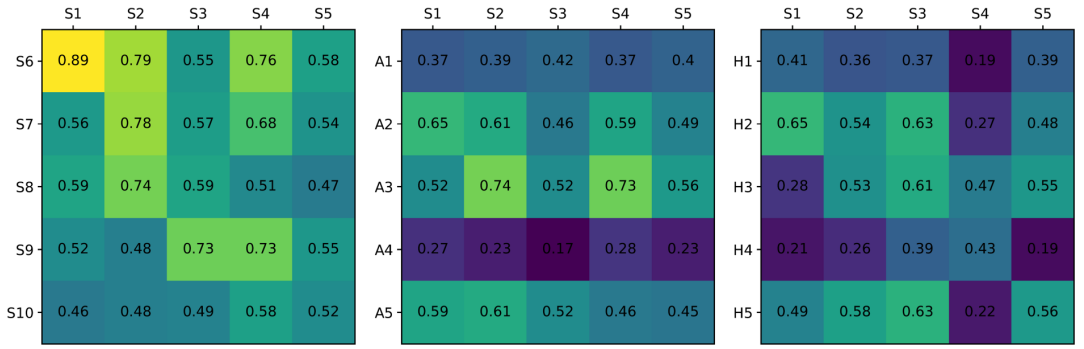
Each element in the correlation matrix is the average value of Pearson's correlation coefficient between every two samples from two subjects. Fig. 18 shows that the correlation of the relation-aware features between the subjects with the same disease labels is more than 14% higher than those with different diseases. After passing the shared module of the RiskNet, the obtained physiological features become more correlated between the same disease and more uncorrelated between different diseases. As observed in Fig. 19, 85% of feature correlation results are higher than 0.5 in the same disease group, whereas only 26% of feature correlation results are higher than 0.5 in the "Stroke VS Alzheimer" group and the "Stroke VS Health" group. Therefore, our developed generic gait feature extraction approach is effective to characterize a particular type of neural disease.

8.2 Compute Efficiency Analysis

In this part, we compare the training efficiency of our approach (i.e., a generic model for multiple diseases) with a conventional approach (i.e., a particular model for a single disease) when introducing a new disease.

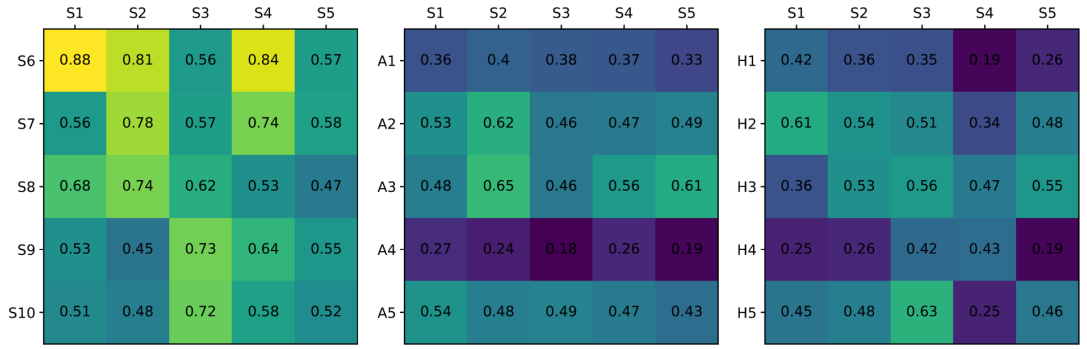
As aforementioned, to implement stroke prediction, our approach freezes the parameters in the gait segment relationship extractor and the shared module of RiskNet, and only opens the customized module of the RiskNet for optimization. In contrast, a conventional approach often trains a single disease detection model from the very beginning. We use the ResNet with 18 layers as our baseline, which is a conventional model for disease detection [40]. For a fair comparison, the training and test datasets are the same for both approaches.

We first compare the training time for each model. The training time consists of the data processing time and epoch time. The data processing of the baseline model includes gait cycle segmentation and spectrogram transformation, which takes about 840 seconds. With the spectrograms as input, each epoch for training the baseline model is around 585 seconds. On the other side, the data processing part of our model is around 1200 seconds, because the obtained spectrograms need to be further fed to our designed gait segment relationship



(a) Stroke VS Stroke (mean corr=0.61) (b) Stroke VS Alzheimer (mean corr=0.47) (c) Stroke VS Health (mean corr=0.43)

Fig. 18. The correlation analysis of global relation-aware features between different groups of subjects.



(a) Stroke VS Stroke (mean corr=0.63) (b) Stroke VS Alzheimer (mean corr=0.43) (c) Stroke VS Health (mean corr=0.41)

Fig. 19. The correlation analysis of shared physiological features between different groups of subjects.

extractor and the shared physiological knowledge learning module successively to obtain generic features before training. Although the data processing takes a bit longer, each epoch time of our model is largely reduced, which only needs 34 seconds. Overall, our model training is 5-10 times faster than the baseline model training depending on the balance between epoch numbers selection and accuracy.

Fig. 20 visualizes the training and validation loss between our model and baseline model. In the baseline model, the validation loss drops slowly, and the overfitting occurs at about the 25th epoch. In contrast, our model has a larger decrease in validation loss, and the overfitting is not obvious. For further exploration, we investigate the influence of the number of training samples on detection performance. As observed in Fig. 21, the F1-score of the baseline model is 4.2% higher than our model for stroke detection with current settings. As the number of training samples is reduced, the F1-score of the baseline model decreases notably to around 57.8%. Different from it, the gap between the performance of our model is very small (less than 4%) regardless of the number of training samples. These results indicate that our model can converge better with relatively small efforts than the baseline model. Therefore, our model is more efficient to update when introducing new diseases with insufficient labeled data.

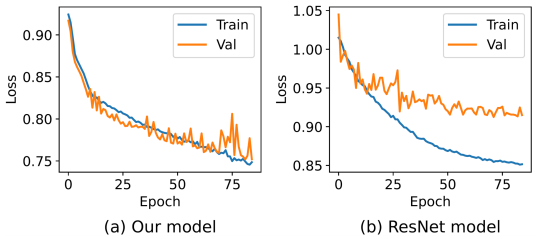


Fig. 20. Train and validation loss comparison.

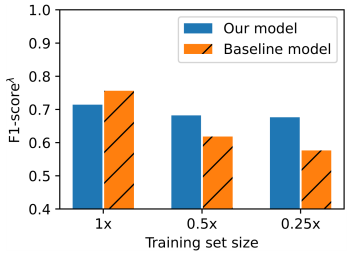


Fig. 21. The impact of training set size on performance.

9 RELATED WORKS

9.1 In-pocket Inertial Sensing

With the advancement of micro-electro-mechanical systems (MEMS) technologies [8], wearable systems have become low-cost, miniature, and easy to be integrated into personal carry-on devices, such as smartphones. In-pocket smartphone sensing has facilitated a set of mobile applications [41–43]. For example, Li et al. [44] leveraged phone inertial sensors to implement indoor localization based on the detection of steps and heading directions. Bayat et al. [45] developed an inertial-based activity recognition system that can identify slow Walking, fast Walking, running, stairs-Up, stairs-Down, and (Aerobic) dancing. Kadhum et al.[46] developed a fall detection system based on smartphone inertial sensors for the elderly.

9.2 Smartphone-based Gait Sensing and Analysis

Smartphone-based gait sensing is mature in step counting, and walking direction detection [47, 48]. Recent studies have revealed the potential of smartphone-based gait sensing and analysis in the healthcare area, including brain disease screening [49], rehabilitation monitoring [50], and medicine effectiveness detection [11]. The gait analysis approaches mainly have two directions. The first direction is to perform signal processing on the inertial data to craft some temporal features (e.g., swing/stance time, single/double support time) and advanced temporal features (e.g., cadence, gait stability, regularity, balance, freeze index) [9, 51]. Since the feature crafting is conducted based on mathematical operations (e.g., average, variance, and RMS), some key pathological information may be lost in this procedure. The second direction is to develop data-driven machine-learning models for feature learning. Patel et al. [12] developed an autoencoder model to extract mobility features from gait inertial data for predicting TBI, but the learned features lack of explainability due to the nature of autoencoder design. Zhang et al. [10] adopted CNN as the gait feature encoder to detect medication adherence of Parkinson's disease. However, such CNN-based feature extractors need to be specialized training for a certain brain disease. Table 3 compares *NeuralGait* with the existing smartphone-based gait analysis system.

9.3 Mobile Health Systems

Mobile health systems have become increasingly popular owing to their convenience and ease of use. SpiroSonic [52] monitored human lung function via acoustic sensing on smartphones. Nandakumar et al. [53] used a smartphone as an active sonar system to emit frequency-modulated acoustic signals and listen to their reflections for detecting sleep apnea events. StressSense [54] unobtrusively recognized the stress from human voice using smartphones embedded microphones. PhO_2 [55] leveraged camera and flashlight functions that are available on smartphones to measure the blood oxygen level. MobileDeepPill [56] used a smartphone built-in camera to recognize unconstrained pill images for medicine management. PDMove [10] used the smartphone's built-in accelerometer and gyroscope to capture the gait change for PD medication adherence detection. PDLens [11]

Table 6. The comparison of smartphone-based gait analysis approaches for brain disease screening.

System	Gait Feature Type	Feature Benefits			Brain Disease Screening Type	Applicable to Unseen Users
		Generic	Explainable	High-compacity		
Howell et al. [51]	Temporal	✓	✓		TBI	
Patel et al. [12]	autoencoder-based	✓		✓	TBI	
PDMove [10]	CNN-based		✓	✓	Medicine adherence detection of PD	
PDLens [11]	CNN-based		✓	✓	Medicine effectiveness detection of PD	✓
NeuralGait	**Relation-based**	✓	✓	✓	**multi-lable PD, TBI, Stroke**	✓

further leveraged both inertial sensors and microphones to collect PD biomarkers to augment drug effectiveness detection.

10 CONCLUSION

In this paper, we introduce *NeuralGait*, a smartphone-based gait sensing and analysis system to facilitate multiple brain disease screening. Specifically, we first leverage an in-pocket smartphone to collect users' gait data in daily life and then develop and prune the gait segments relationship encoding model to extract the time-relation feature, frequency-relation feature, and global relation-aware feature. After that, these gait features are transmitted to a scalable RiskNet in the cloud for estimating multiple brain disease risks. We further adopt a threshold-based voting method for a more reliable prediction of each disease. A pathological hot-update protocol is applied to efficiently add new brain diseases in the RiskNet to maximize the system's applicability. Extensive experiments indicate that the *NeuralGait* can achieve above 76.1% sensitivity of screening PD and stroke, and a 5.9% false alarm rate for screening TBI.

ACKNOWLEDGMENTS

We thank all anonymous reviewers for their insightful comments on this paper. This work was supported by the National Science Foundation under grants OISE-2106996 and CNS-2050910.

REFERENCES

[1] Y. Wang, Y. Pan, and H. Li, "What is brain health and why is it important?" *BMJ*, vol. 371, 2020. [Online]. Available: https://www.bmj.com/content/371/bmj.m3683
[2] C. Marras, J. Beck, J. Bower, E. Roberts, B. Ritz, G. Ross, R. Abbott, R. Savica, S. Van Den Eeden, A. Willis *et al.*, "Parkinson's foundation p4 group," *Prevalence of Parkinson's disease across North America. NPJ Parkinsons Dis*, vol. 4, no. 21, p. 21, 2018.
[3] "Dementia." [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/dementia
[4] S. S. Virani, A. Alonso, H. J. Aparicio, E. J. Benjamin, M. S. Bittencourt, C. W. Callaway, A. P. Carson, A. M. Chamberlain, S. Cheng, F. N. Delling *et al.*, "Heart disease and stroke statistics—2021 update: a report from the american heart association," *Circulation*, vol. 143, no. 8, pp. e254–e743, 2021.
[5] "Traumatic brain injury / concussion," Dec 2021. [Online]. Available: https://www.cdc.gov/TraumaticBrainInjury/
[6] A. M. Wennberg, R. Savica, and M. M. Mielke, "Association between various brain pathologies and gait disturbance," *Dementia and geriatric cognitive disorders*, vol. 43, no. 3-4, pp. 128–143, 2017.
[7] A. H. Snijders, B. P. Van De Warrenburg, N. Giladi, and B. R. Bloem, "Neurological gait disorders in elderly people: clinical approach and classification," *The Lancet Neurology*, vol. 6, no. 1, pp. 63–74, 2007.
[8] A. Filippeschi, N. Schmitz, M. Miezal, G. Bleser, E. Ruffaldi, and D. Stricker, "Survey of motion tracking methods based on inertial sensors: A focus on upper limb human motion," *Sensors*, vol. 17, no. 6, p. 1257, 2017.
[9] S. Chen, J. Lach, B. Lo, and G.-Z. Yang, "Toward pervasive gait analysis with wearable sensors: A systematic review," *IEEE journal of biomedical and health informatics*, vol. 20, no. 6, pp. 1521–1537, 2016.

[10] H. Zhang, C. Xu, H. Li, A. S. Rathore, C. Song, Z. Yan, D. Li, F. Lin, K. Wang, and W. Xu, "Pdmove: Towards passive medication adherence monitoring of parkinson's disease using smartphone-based gait assessment," *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies*, vol. 3, no. 3, pp. 1–23, 2019.
[11] H. Zhang, G. Guo, C. Song, C. Xu, K. Cheung, J. Alexis, H. Li, D. Li, K. Wang, and W. Xu, "Pdlens: smartphone knows drug effectiveness among parkinson's via daily-life activity fusion," in *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*, 2020, pp. 1–14.
[12] B. Patel, S. Srikanthan, F. Asanit, and E. Agu, "Machine learning prediction of tbi from mobility, gait and balance patterns," in *2021 IEEE/ACM Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*. IEEE, 2021, pp. 11–22.
[13] E. Yiou, T. Caderby, A. Delafontaine, P. Fourcade, and J.-L. Honeine, "Balance control during gait initiation: State-of-the-art and research perspectives," *World journal of orthopedics*, vol. 8, no. 11, p. 815, 2017.
[14] W. Yin, K. Kann, M. Yu, and H. Schütze, "Comparative study of cnn and rnn for natural language processing," *arXiv preprint arXiv:1702.01923*, 2017.
[15] A. L. Rosso, S. A. Studenski, W. G. Chen, H. J. Aizenstein, N. B. Alexander, D. A. Bennett, S. E. Black, R. Camicioli, M. C. Carlson, L. Ferrucci *et al.*, "Aging, the central nervous system, and mobility," *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*, vol. 68, no. 11, pp. 1379–1386, 2013.
[16] K. Takakusaki, "Neurophysiology of gait: from the spinal cord to the frontal lobe," *Movement Disorders*, vol. 28, no. 11, pp. 1483–1491, 2013.
[17] I. Pisotta and M. Molinari, "Cerebellar contribution to feedforward control of locomotion," *Frontiers in human neuroscience*, vol. 8, p. 475, 2014.
[18] A. C. LeBlanc, "The role of apoptotic pathways in alzheimer's disease neurodegeneration and cell death," *Current Alzheimer Research*, vol. 2, no. 4, pp. 389–402, 2005.
[19] R. Iansek, F. Huxham, and J. McGinley, "The sequence effect and gait festination in parkinson disease: contributors to freezing of gait?" *Movement disorders: official journal of the Movement Disorder Society*, vol. 21, no. 9, pp. 1419–1424, 2006.
[20] J. Ghajar, "Traumatic brain injury," *The Lancet*, vol. 356, no. 9233, pp. 923–929, 2000.
[21] P. C. Fino, L. Parrington, W. Pitt, D. N. Martini, J. C. Chesnutt, L.-S. Chou, and L. A. King, "Detecting gait abnormalities after concussion or mild traumatic brain injury: a systematic review of single-task, dual-task, and complex gait," *Gait & Posture*, vol. 62, pp. 157–166, 2018.
[22] D. N. Martini, M. J. Sabin, S. A. DePesa, E. W. Leal, T. N. Negrete, J. J. Sosnoff, and S. P. Broglio, "The chronic effects of concussion on gait," *Archives of physical medicine and rehabilitation*, vol. 92, no. 4, pp. 585–589, 2011.
[23] T. Fonseca, P. Cortes, J. Monteiro, V. Salgado, J. Ferro, A. Franco, J. Nogueira, and J. da Costa, "Acute cerebrovascular disorder and arterial hypertension. prospective study with 248 patients," *Revista portuguesa de cardiologia: orgao oficial da Sociedade Portuguesa de Cardiologia= Portuguese journal of cardiology: an official journal of the Portuguese Society of Cardiology*, vol. 15, no. 7-8, pp. 565–73, 1996.
[24] K. F. de Laat, A. G. van Norden, R. A. Gons, L. J. van Oudheusden, I. W. van Uden, B. R. Bloem, M. P. Zwiers, and F.-E. de Leeuw, "Gait in elderly with cerebral small vessel disease," *Stroke*, vol. 41, no. 8, pp. 1652–1658, 2010.
[25] A. D. Kuo, "The six determinants of gait and the inverted pendulum analogy: A dynamic walking perspective," *Human movement science*, vol. 26, no. 4, pp. 617–656, 2007.
[26] A. Behboodi, H. Wright, N. Zahradka, and S. C. Lee, "Seven phases of gait detected in real-time using shank attached gyroscopes," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2015, pp. 5529–5532.
[27] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly *et al.*, "An image is worth 16x16 words: Transformers for image recognition at scale," *arXiv preprint arXiv:2010.11929*, 2020.
[28] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in *Advances in neural information processing systems*, 2017, pp. 5998–6008.
[29] P. Michel, O. Levy, and G. Neubig, "Are sixteen heads really better than one?" *arXiv preprint arXiv:1905.10650*, 2019.
[30] W. F. Pickard, "Generalizations of the goldman-hodgkin-katz equation," *Mathematical biosciences*, vol. 30, no. 1-2, pp. 99–111, 1976.
[31] J. C. Eccles, *The physiology of synapses*. Academic Press, 2013.
[32] G. D. Lamb, "Excitation-contraction coupling in skeletal muscle: comparisons with cardiac muscle." *Clinical and experimental pharmacology & physiology*, vol. 27, no. 3, pp. 216–224, 2000.
[33] C.-L. Zhang, J.-H. Luo, X.-S. Wei, and J. Wu, "In defense of fully connected layers in visual representation transfer," in *Pacific Rim Conference on Multimedia*. Springer, 2017, pp. 807–817.
[34] M. S. Sorower, "A literature survey on algorithms for multi-label learning," *Oregon State University, Corvallis*, vol. 18, pp. 1–25, 2010.
[35] S. Godbole and S. Sarawagi, "Discriminative methods for multi-labeled classification," in *Pacific-Asia conference on knowledge discovery and data mining*. Springer, 2004, pp. 22–30.
[36] W. Maetzler, I. Liepelt, and D. Berg, "Progression of parkinson's disease in the clinical phase: potential markers," *The Lancet Neurology*, vol. 8, no. 12, pp. 1158–1171, 2009.
[37] G. A. Kang, J. M. Bronstein, D. L. Masterman, M. Redelings, J. A. Crum, and B. Ritz, "Clinical characteristics in early parkinson's disease in a central california population-based study," *Movement disorders: official journal of the Movement Disorder Society*, vol. 20, no. 9, pp.

1133–1142, 2005.
[38] S.-u. Ko, M. I. Tolea, J. M. Hausdorff, and L. Ferrucci, "Sex-specific differences in gait patterns of healthy older adults: results from the baltimore longitudinal study of aging," *Journal of biomechanics*, vol. 44, no. 10, pp. 1974–1979, 2011.
[39] S.-u. Ko, S. Stenholm, E. J. Metter, and L. Ferrucci, "Age-associated gait patterns and the role of lower extremity strength–results from the baltimore longitudinal study of aging," *Archives of gerontology and geriatrics*, vol. 55, no. 2, pp. 474–479, 2012.
[40] M. Talo, O. Yildirim, U. B. Baloglu, G. Aydin, and U. R. Acharya, "Convolutional neural networks for multi-class brain disease detection using mri images," *Computerized Medical Imaging and Graphics*, vol. 78, p. 101673, 2019.
[41] J. Qian, J. Ma, R. Ying, P. Liu, and L. Pei, "An improved indoor localization method using smartphone inertial sensors," in *International Conference on Indoor Positioning and Indoor Navigation*. IEEE, 2013, pp. 1–7.
[42] J. Shang, F. Gu, X. Hu, and A. Kealy, "Apfiloc: An infrastructure-free indoor localization method fusing smartphone inertial sensors, landmarks and map information," *Sensors*, vol. 15, no. 10, pp. 27 251–27 272, 2015.
[43] M. Shoaib, H. Scholten, and P. J. Havinga, "Towards physical activity recognition using smartphone sensors," in *2013 IEEE 10th international conference on ubiquitous intelligence and computing and 2013 IEEE 10th international conference on autonomic and trusted computing*. IEEE, 2013, pp. 80–87.
[44] F. Li, C. Zhao, G. Ding, J. Gong, C. Liu, and F. Zhao, "A reliable and accurate indoor localization method using phone inertial sensors," in *Proceedings of the 2012 ACM conference on ubiquitous computing*, 2012, pp. 421–430.
[45] A. Bayat, M. Pomplun, and D. A. Tran, "A study on human activity recognition using accelerometer data from smartphones," *Procedia Computer Science*, vol. 34, pp. 450–457, 2014.
[46] A. Kadhum, H. Al-Libawy, and E. Hussein, "An accurate fall detection system for the elderly people using smartphone inertial sensors," in *Journal of Physics: Conference Series*, vol. 1530, no. 1. IOP Publishing, 2020, p. 012102.
[47] A. Brajdic and R. Harle, "Walk detection and step counting on unconstrained smartphones," in *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*, 2013, pp. 225–234.
[48] N. Roy, H. Wang, and R. Roy Choudhury, "I am a smartphone and i can tell my user's walking direction," in *Proceedings of the 12th annual international conference on Mobile systems, applications, and services*, 2014, pp. 329–342.
[49] R. J. Ellis, Y. S. Ng, S. Zhu, D. M. Tan, B. Anderson, G. Schlaug, and Y. Wang, "A validated smartphone-based assessment of gait and gait variability in parkinson's disease," *PLoS one*, vol. 10, no. 10, p. e0141694, 2015.
[50] T.-V. How, J. Chee, E. Wan, and A. Mihailidis, "Mywalk: a mobile app for gait asymmetry rehabilitation in the community," in *2013 7th International Conference on Pervasive Computing Technologies for Healthcare and Workshops*. IEEE, 2013, pp. 73–76.
[51] D. R. Howell, V. Lugade, M. Taksir, and W. P. Meehan III, "Determining the utility of a smartphone-based gait evaluation for possible use in concussion management," *The Physician and sportsmedicine*, vol. 48, no. 1, pp. 75–80, 2020.
[52] X. Song, B. Yang, G. Yang, R. Chen, E. Forno, W. Chen, and W. Gao, "Spirosonic: monitoring human lung function via acoustic sensing on commodity smartphones," in *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*, 2020, pp. 1–14.
[53] R. Nandakumar, S. Gollakota, and N. Watson, "Contactless sleep apnea detection on smartphones," in *Proceedings of the 13th annual international conference on mobile systems, applications, and services*, 2015, pp. 45–57.
[54] H. Lu, D. Frauendorfer, M. Rabbi, M. S. Mast, G. T. Chittaranjan, A. T. Campbell, D. Gatica-Perez, and T. Choudhury, "Stresssense: Detecting stress in unconstrained acoustic environments using smartphones," in *Proceedings of the 2012 ACM conference on ubiquitous computing*, 2012, pp. 351–360.
[55] N. Bui, A. Nguyen, P. Nguyen, H. Truong, A. Ashok, T. Dinh, R. Deterding, and T. Vu, "Pho2: Smartphone based blood oxygen level measurement systems using near-ir and red wave-guided light," in *Proceedings of the 15th ACM conference on embedded network sensor systems*, 2017, pp. 1–14.
[56] X. Zeng, K. Cao, and M. Zhang, "Mobiledeeppill: A small-footprint mobile deep learning system for recognizing unconstrained pill images," in *Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services*, 2017, pp. 56–67.