A Hybrid Method for Implicit Intention Inference Based on Punished-Weighted Naïve Bayes

Zheng Gao, Shiqian Wu, *Senior Member, IEEE*, Zhonghua Wan, and Sos Agaian, *Fellow, IEEE*

***Abstract*— Gaze-based implicit intention inference provides a new human-robot interaction for people with disabilities to accomplish activities of daily living independently. Existing gaze-based intention inference is mainly implemented by the data-driven method without prior object information in intention expression, which yields low inference accuracy. Aiming to improve the inference accuracy, we propose a gaze-based hybrid method by integrating model-driven and data-driven intention inference tailored to disability applications. Specifically, intention is considered as the combination of verbs and nouns. The objects corresponding to the nouns are regarded as intention-interpreting objects and served as prior knowledge, i.e., punished factors. The punished factor considers the object information, i.e., the priority in object selection. Class-specific attribute weighted naïve Bayes model learned through training data is presented to represent the relationship among intentions and objects. An intention inference engine is developed by combining the human prior knowledge, and the data-driven class-specific attribute weighted naïve Bayes model. Computer simulations: (i) verify the contribution of each critical component of the proposed model, (ii) evaluate the inference accuracy of the proposed model, and (iii) show that the proposed method is superior to state-of-the-art intention inference methods in terms of accuracy.**

***Index Terms*— Intention inference, intention-interpreting object, prior knowledge model, punished factor, class-specific attribute weighted naïve bayes.**

I. Introduction

WITH the increasing number of people with weak functional groups such as amyotrophic lateral sclerosis (ALS), how to help the groups with limited motion to complete activities of daily living has attracted widespread attention from business to academia [1]. Current works focus on designing human-robot interactions to build auxiliary robots to help the weak functional groups accomplish daily life activities [2], [3]. The critical issue in designing auxiliary robots is to infer human intentions via simple and efficient human-computer interaction methods. At present, a variety of human-computer interaction technologies have been widely used in intention inference, such as speech [4], EEG [5], EMG [6], and so on.

Visual attention signal is another way of human-computer interaction [7], [8], [9]. Studies have shown that human visual attention can express the mental processes revealed by gaze cues, which are associated with visual objects [10], [11], [12]. Previous work has studied eye-tracking technology to estimate the gaze point in the scene [13] and designed an auxiliary robot to help weak functional groups such as ALS [14]. These patients do not have full abilities of muscle movements but still retain the ability to control eye movement. Therefore, it is feasible to infer the intentions of the disabled through the human vision to improve their quality of life.

Human implicit intention inference refers to the procedure: 1) capturing human eye movement through eye tracking devices, 2) identifying the objects of interest, and 3) inferring the users' intention. An example is shown in Fig.1. When a person wants to drink coffee, he/she gazes at the coffee, cup, and kettle. Then we infer that he/she may want to "drink coffee." Existing works capture eye movements and identify objects of interest [15], [16]. In this paper, we focus on how to infer human intention through the objects of interest (Step 3).

Previous intention inference works mainly adopted machine learning methods [15], [17], which require training data to learn the classification models. These approaches extract the mapping relationship between objects and intentions from the intention database. Besides, the model parameters are

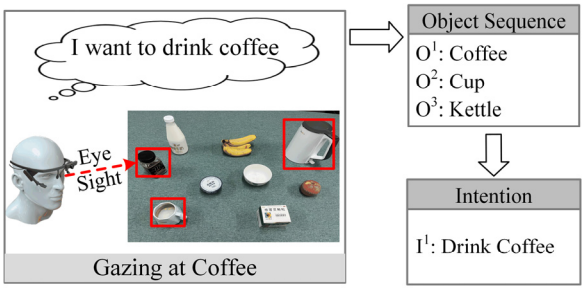


Fig. 1. An example of intention inference, i.e., the intention is inferred by the objects: coffee, cup, and kettle at which a person gazes.

Manuscript received 23 September 2022; revised 4 January 2023 and 26 February 2023; accepted 15 March 2023. Date of publication 24 March 2023; date of current version 29 March 2023. This work was supported in part by the New Initiative Project under Grant 2018TDX06, Wuhan University of Science and Technology. *(Corresponding author: Shiqian Wu.)*

Zheng Gao, Shiqian Wu, and Zhonghua Wan are with the School of Information Science and Engineering, Institute of Robotics and Intelligent Systems, Wuhan University of Science and Technology, Wuhan 430081, China (e-mail: gzheng@wust.edu.cn; shiqian.wu@wust.edu.cn; zhwan@wust.edu.cn).

Sos Agaian is with the Department of Computer Science, College of Staten Island and Graduate Center, City University of New York, New York City, NY 10314 USA (e-mail: sos.agaian@csi.cuny.edu).

Digital Object Identifier 10.1109/TNSRE.2023.3259550

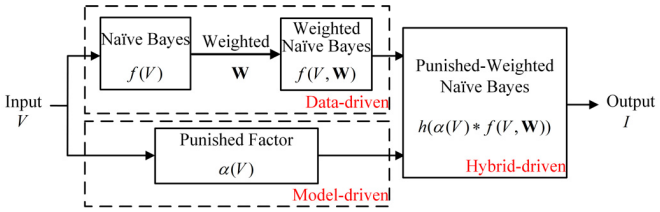


Fig. 2. The flowchart of the proposed hybrid-driven intention inference method. V is the visual object sequence and I is the inferred intention.

determined by the inference performance on training data. It is noted that the object selection for intention expression exists prior pattern, i.e., the objects associated with the intention have different priorities. The above data-driven methods view each object equally in model training and ignore the priority of object selection [18].

In this paper, we utilize the prior knowledge of object priority and propose a combined data-driven and human prior knowledge model-driven approach to accomplish intention inference, as shown in Fig.2. A class-specific attribute weighted naïve Bayes (CAWNB) model is introduced to compensate for the lack of data expressiveness due to the conditional independence assumption of object selection in the naïve Bayes. We refer to the objects with the highest priority as intention-interpreting objects, which are served as prior knowledge, i.e., punished factors.

The proposed punished-weighted naïve Bayes (P-WNB), a hybrid model by combining the data-driven weighted naïve Bayes model and the model-driven intention with prior knowledge, performs the intention inference work. The main contributions are summarized as follows:

- A weighted naïve Bayes model is proposed to increase the weights of highly predictive objects and weaken object selection's conditional independence during intention representation.
- The intention-interpreting object is served as a punished factor and adopted into intention inference.
- We demonstrate a) the advantage of the priority of the intention-interpreting object and b) the effectiveness of the proposed algorithm in comparison with the state-of-the-art methods.

The remainder of this paper is organized as follows. Section II analyzes related work on gaze-based intention inference and weighted naïve Bayes. Section III describes our model for intention inference. Experimental results are given in Section IV. We discuss the advantages and limitations of the proposed model in Section V, followed by the conclusion drawn in Section VI.

II. Related Work

A. Gaze-Based Intention Inference

Gaze-based implicit intention inference is to infer human intention through the attentive object sequence. Existing works adopted data-driven machine learning methods, such as support vector machine (SVM) [15], [17], naïve Bayes [18], Decision Tree [19], etc. Li et al. [17] first treated each object as one feature and built an SVM classifier to infer several daily intentions, such as "Prepare a cup of coffee," "Prepare oatmeal for breakfast," and so on. The same approach was also used to predict which ingredient the customers intended to request in a dyadic sandwich-making scenario [20]. However, the dependence on the amount of training data makes this less usable. To this end, Li et al. [18] extended their previous work to build a naïve Bayes model which performs better with a small amount of training data to represent the corresponding relationship between intentions and objects.

The intention with maximum posterior probability is the final inference result. Experimental results showed that there was a priority among the objects when expressing a specific intention, but the priority was not reflected in the intention knowledge base and inference machine.

Another prominent work was proposed by Koochaki et al. [19]. After acquiring the region of interest (ROI) by clustering gaze points through Density-Based Spatial Clustering of Applications with Noise (DBSCAN), the object semantic information in ROI by convolution neural networks (CNN) is identified, and intention inference is accomplished by machine learning methods, such as SVM [15], random forest, decision tree, and Adaboost. The pre-trained hidden Markov model (HMM) was adopted to identify the selection order of ROIs, and the CNN extracted the high-level features of each ROI, which were then fed into long short-term memory (LSTM) to learn the temporal dependency between these features [16]. The advantage of this method over prior works [15], [17], [18] is that the temporal dependence between different objects is taken into account.

B. Weighted Naïve Bayes

The Naive Bayes learning scheme is well-used on most classification tasks because of its simplicity, effectiveness, interpretability, and computational efficiency. However, this approach makes a key assumption that all attributes are entirely independent of each other given class, which is rarely accurate in real-world applications and would harm its performance with attribute dependencies. Recent studies have shown that attribute weighting approaches can improve naive Bayes with attribute redundancy. Attribute weighted is a flexible way to alleviate the conditional independence assumption [21]. Existing attribute weighting methods can be mainly divided into two categories [22]: 1) filter approaches and 2) wrapper approaches.

Filter approaches use the heuristic information of training data to calculate the weights of the attributes. The decision tree-based feature weighting method assigns the weight to each attribute through the average minimum depth of 10 unpruned decision trees [23]. Jiang et al. [24] proposed a correlation-based feature weighting method, and the weight is proportional to the difference between the feature-class correlation and the average feature-feature intercorrelation. Some other methods determine the feature weights through the attribute gain ratio [25], the Kullback-Leibler measure [26], and so on.

TABLE I
PROS AND CONS OF 4 TYPICAL WEIGHTED NAÏVE BAYES

	Bayes Weight Definitions	Advantages	Limitations
Jiang *et al.* [24]	Correlation-Based Feature Weighting Naive Bayes (CFWNB) $I = \arg\max_{I_i \in \{I_1, I_2, \ldots, I_N\}} P(I_i) \prod_{j=1}^{M} P(O_j = o_j \mid I_i)^{w_j}$	The highly predictive features are highly correlated with the class yet uncorrelated with other features.	Differences in the importance of each attribute to each class are not considered.
Jiang *et al.* [28]	Class-specific attribute weighted naïve Bayes (CAWNB) $I = \arg\max_{I_i \in \{I_1, I_2, \ldots, I_N\}} P(I_i) \prod_{j=1}^{M} P(O_j = o_j \mid I_i)^{w_{i,j}}$	The attribute importance for each class should be different is considered.	Assigning weights by raising the power of the conditional probabilities may cause conditional probabilities to behave inversely.
Foo *et al.* [29]	Attribute weighted naïve bayes (AWNB) $I = \arg\max_{I_i \in \{I_1, I_2, \ldots, I_N\}} P(I_i) \prod_{j=1}^{M} P(O_j = o_j \mid I_i)^{(\exp(-w_j))}$	Predictive attributes are assigned higher weights and avoid conditional probabilities of behaving inversely.	Differences in the importance of each attribute to each class are not considered.
Frank *et al.* [30]	Locally Weighted Naïve Bayes (LWNB) $P(I_i) = \frac{1 + \sum_{k=1}^{K} I(I_k = I_i) w_k}{N + \sum_{k=1}^{K} w_k}$, $P(O_j = o_j \mid I_i) = \frac{1 + \sum_{k=1}^{K} I(O_j = o_j) I(I_k = I_i) w_k}{n_j + \sum_{k=1}^{K} I(O^i = o_j) w_k}$, $I = \arg\max_{I_i \in \{I_1, I_2, \ldots, I_N\}} P(I_i) \prod_{j=1}^{M} P(O_j = o_j \mid I_i)$	Used as linear regression, it is easy to fit into the training data.	The effort involved in learning is deferred to classification time, which leads to slow inference.

Note: CFWNB, CAWNB, and AWNB are attributed weighting methods,whereas LWNB is an instance selection method.

Wrapper approaches use the classification performance on training data to determine attribute weights. The weights are learned from training data by minimizing the classification error. Generally, the mean squared error and the conditional log-likelihood are used as objective functions. For example, the differential evolution algorithm is adopted to learn the optimal attribute weights [27] by assigning each attribute the same weight for all classes. Recently in [28], Jiang et al. (i) argue that NB attribute weighting should be class-specific (class- dependent) i.e., the importance of each attribute for each class should differ, (ii) present a new class-specific attribute weighting model (CAWNB), and (ii) propose two gradient-based learning algorithms using L-BFGS-M method [31] an optimization in the family of the limited-memory quasi-Newton method, to learn such an attribute weight matrix. In this work, we focused on applying the weighted naïve Bayes in intention inference [21]. Table I summarizes four representative weighted naïve Bayes schemes.

III. INTENTION INFERENCE BASED ON P-WNB

Human eye motion behavior is related to visual objects, and multiple objects are involved in complex behaviors. Thus, human intention I can be inferred by the perceived object sequence $(O^1, O^2, \ldots, O^T)$. I is expressed as

$$I = f(O^1, O^2, \ldots, O^T), \tag{1}$$

and

$$O^t \in \{O_1, O_2, \ldots, O_M\}, \tag{2}$$

where T represents the number of perceived objects, $t \in \{1, 2, \ldots, T\}$, and M indicates the total types of objects, i.e., the number of features.

The proposed punished-weighted naïve Bayes (P-WNB) for intention inference can be expressed as follows:

$$I = \arg\max_{I_i \in \{I_1, I_2, \ldots, I_N\}} \alpha(I_i | O^1, O^2, \ldots, O^T) P(I_i) \prod_{j=1}^{M} P(O_j = o_j | I_i)^{w_{i,j}}, \tag{3}$$

where N is the number of possible intentions, $\alpha(I_i | O^1, O^2, \ldots, O^T)$ is the punished factor, and $w_{i,j}$ is the weight in class-specific attribute weighted naïve Bayes of the j-th attribute object O_j for class I_i, $P(I_i)$ and $P(O_j = o_j | I_i)$ are obtained by statistics in experiments and stored in the database of naïve Bayes, and

$$o_j = \begin{cases} 1, & \text{if } O_j \text{ is viewed} \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

Generally, the naïve Bayesian learning is robust on the small amount of noise. The performance of naïve Bayesian learning is limited due to the hypothesis that all features are equally important and independent (conditional) given the class value, which usually does not hold in a real-world scenario. Several methodologies have been offered to reduce the element of the independence assumption. For example, assign proper weight to each feature by estimating how important the feature is, give different attributes with a higher weight for attributes that have more influences, or discriminatively assign each attribute a specific weight for each class, which different weights are. The critical problem of the weighting approach is how to assign each attribute the weight for each class and how to process the weight learning.

CAWNB is appropriate for intention inference since each object's importance for each intention is different, and the human-computer interaction system has a higher demand for inference speed than training speed.

A. Data-Driven Class-Specific Attribute Weighted Naïve Bayes

Given a test object sequence $(O^1, O^2, \ldots, O^T)$, naïve Bayes uses (5) to predict its intention label.

$$I = \underset{I_i \in \{I_1, I_2, \ldots, I_N\}}{\arg\max} P(I_i) \prod_{j=1}^{M} P(O_j = o_j | I_i). \tag{5}$$

CAWNB assigns a unique weight to each feature with each class [28]. Then the intention is predicted by

$$I = \underset{I_i \in \{I_1, I_2, \ldots, I_N\}}{\arg\max} P(I_i) \prod_{j=1}^{M} P(O_j = o_j | I_i)^{w_{i,j}}, \tag{6}$$

where $w_{i,j}$ indicates the weight between i-th intention and j-th object. The weight matrix can be expressed as

$$\begin{matrix} I_1 \rightarrow \\ - \\ I_N \rightarrow \end{matrix} \begin{bmatrix} w_{1,1} & \cdots & w_{1,M} \\ \vdots & \ddots & \vdots \\ w_{N,1} & \cdots & w_{N,M} \end{bmatrix}_{N \times M} = \mathbf{W} \qquad \begin{matrix} \uparrow & & \uparrow \\ O_1 & - & O_M \end{matrix} \tag{7}$$

which illustrates the class-specific attribute weights.

To learn this weight matrix, a data-driven learning approach is adopted. Each element in $\mathbf{W}$ is initialized to 1, which implies that CAWNB is learned from the original naïve Bayes. Thus, the problem of the weight learning process is transferred to the optimization problem of the objective function. The optimization function is given the conditional log-likelihood function (CLL). The optimization function is

$$\begin{aligned} \max_{w \in [0,1]} CLL(w) &= \sum_{m=1}^{|D|} \log \hat{P}(I^m | \mathbf{O}^m; \mathbf{W}) \\ &= \sum_{m=1}^{|D|} \log \frac{\gamma_{I_i O}(\mathbf{W})}{\sum_{I_i} \gamma_{I_i O}(\mathbf{W})}, \end{aligned} \tag{8}$$

where $|D|$ is the training data, and

$$\gamma_{I_i O}(\mathbf{W}) = P(I_i) \prod_{j=1}^{M} P(O_j = o_j | I_i)^{w_{i,j}}. \tag{9}$$

In this work, 70% of data collected by experiments is used to train this model. We use the JADE [32] optimization procedure to solve this optimization problem. It is noted that JADE is used to minimize the objective function, whereas this work is to maximize the conditional log-likelihood function. Therefore, the opposite of the conditional log-likelihood function is taken when using JADE to optimize the solution.

CAWNB assigns higher weights to more predictive objects to weaken object selection's conditional independence. This allows the inference machine to be more discriminative. Refer to [28] for more details of CAWNB.

B. Model-Driven Punished Factor

1) Priority of Object Selection: Data-driven CAWNB still ignores the priorities in object selection. To better understand the priority of objects, we first define three terms: intention-associated objects, dominant objects, and intention-interpreting objects. The intention-associated objects are the objects that may be viewed in intention expression. The dominant objects are the objects with a high probability of being viewed in intention inference. The identification of the dominant objects is obtained by analyzing experimental statistics. Each intention has a unique set of dominant objects, but each dominant object still has a different priority [18]. The dominant object with the highest priority is defined as the intention-interpreting object.

An intention can usually be a combination of verbs and nouns [33]. The verb indicates the action to be taken, and the nouns indicate the operational objects. For implicit intention inference, the operational object reflects the preferred object among intention-associated objects. In this study, the objects corresponding to the nouns are regarded as intention-interpreting objects. For example, the intention-interpreting object of "Drink Milk" is milk, while the intention-interpreting objects of "Drink A Cup of Milk" are cup and milk. An intention may have multiple intention-interpreting objects and the intention-interpreting objects have the identical highest priority, which is validated in Section IV-B.

2) Priority Modeling: For a specific intention, the number of objects corresponding to the nouns is finite, which indicates that the number of objects with the highest priority is finite. The intention-interpreting objects will be noticed as much as possible, even if these objects are not the first to be selected due to environmental effects [18]. Thus, this pattern is modeled as the satisfaction of the intention-interpreting objects of each intention, i.e., the percentage of selected objects among intention-interpreting objects. It can be denoted as

$$\alpha(I_i | O^1, O^2, \ldots, O^t) = \frac{num(I_i | O^1, O^2, \ldots, O^T) + \frac{1}{N}}{num(I_i)}, \tag{10}$$

where $\alpha(I_i | O^1, O^2, \ldots, O^T)$ is the proposed punished factor, $num(I_i)$ is the total number of intention-interpreting objects for intention I_i, $num(I_i | O^1, O^2, \ldots, O^T)$ indicates the number of the intention-interpreting objects in visual sequence $(O^1, O^2, \ldots, O^T)$ for I_i, and $1/N$ is the smoothing term. Its significance is that when all observed objects are not any intention-interpreting object for all intentions, it makes intention inference through this non-intention-interpreting object without making the posterior probability of all intentions to be 0.

From (10), it can be concluded that $\alpha(I_i | O^1, O^2, \ldots, O^T)$ is usually less than 1. α is only slightly greater than 1 when all the intention-interpreting objects of the corresponding intention are selected. This is why we call it the punished factor.

The key problem in getting $num(I_i)$ and $num(I_i | O^1, O^2, \ldots, O^T)$ is how to model intention-interpreting objects. Our idea comes from the representation of the conditional probability matrix $\mathbf{P}$ in naïve Bayes.

P is expressed as

$$\mathrm{P} = \begin{bmatrix} P(O_1|I_1) & \cdots & P(O_M|I_1) \\ \vdots & \ddots & \vdots \\ P(O_1|I_N) & \cdots & P(O_M|I_N) \end{bmatrix}_{N\times M}, \tag{11}$$

where $P(O_j|I_i)$ indicates the strength of correlation between O_j and I_i.

Unified modeling facilitates uniform knowledge representation and computing. Thus, we set up an intention-interpreting object matrix **E**, a prior knowledge matrix, to represent the intention-interpreting object for each intention. **E** is denoted as

$$\begin{array}{c} I_1 \rightarrow \\ - \\ I_N \rightarrow \end{array} \begin{bmatrix} E_{11} & \cdots & E_{1M} \\ \vdots & \ddots & \vdots \\ E_{N1} & \cdots & E_{NM} \end{bmatrix}_{N\times M} = \mathbf{E}, \quad \begin{array}{c} \uparrow \quad \uparrow \\ O_1 \quad O_M \end{array} \tag{12}$$

and

$$E_{i,j} = \begin{cases} 1, & \text{if } O_j \text{ is the intention-interpreting object of } I_i \\ 0, & \text{Otherwise.} \end{cases} \tag{13}$$

The dimension of **E** is the same as **P**. This allows the calculation process to be carried out through matrix operations. Then the total number of intention-interpreting objects for each intention I_i can be obtained from

$$num(I_i) = \sum_{j=1}^{M} E_{i,j}. \tag{14}$$

Besides $num(I_i|O^1, O^2, \ldots, O^T)$ can be obtained directly from **E** and visual sequence $(O^1, O^2, \ldots, O^T)$. Combining CAWNB and the proposed punished factor, our final model P-WNB can be expressed as

$$I = \underset{I_i\in\{I_1,I_2,\ldots,I_N\}}{\arg\max} \alpha(I_i|O^1, O^2, \ldots, O^T)P(I_i) \prod_{j=1}^{M} P(O_j = o_j|I_i)^{w_{i,j}}. \tag{15}$$

Thus, human intention can be inferred by (15). In the next Section, experiments will be set up to verify the reliability of the proposed model.

IV. Experimental Results

This section presents the experimental setup followed by a description of the pattern validation of object selection. Subsequently, we build an intention database, by which we conduct inference experiments and compare our method with existing methods. Additionally, we conduct ablation studies to analyze the benefits of each component and test the proposed method on a challenging dataset.

A. Experimental Setup

Experiments are carried out to record the subjects' attentive objects when expressing intentions. We use the Pupil Core eye tracker to capture the subject's attention. Pupil Core, as described in [34], is an eyeglass-like head-mounted eye tracker, which mainly consists of one world camera, one nose support, and two eye cameras, as shown in Fig.3(a). Fig.3(b) demonstrates the eye tracker's ability to capture both scene and pupil images, where the gaze point coordinates on the test scene can be obtained through the acquired pupil and scene images, as shown in Fig.3(c).

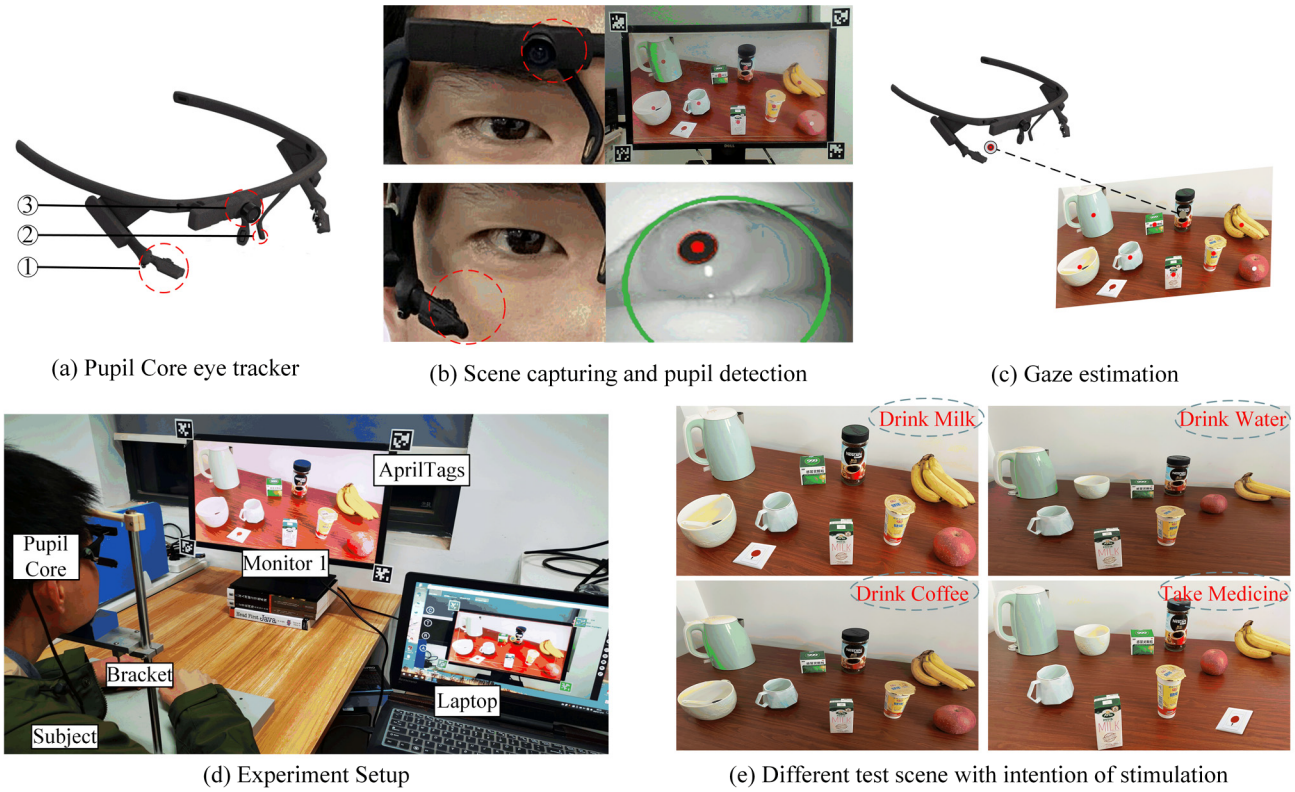
Experiments are carried out in a bright room. The setup included an eye tracker, a laptop for data recording, and a monitor for a scene showing, as shown in Fig.3(d). The participant, wearing the eye tracker, sits in front of the monitor. To ensure eye-tracking accuracy, the height of the subject's head is kept horizontal to the height of the center of monitor 1 as much as possible.

B. Pattern Validation of Object Selection

We present the experimental procedure for validating the pattern of object selection. A special set of intentions is set up and divided into two groups to facilitate the development of comparison experiments. The inclusion of "Take Medicine" is aimed at preventing the object selection preference caused by the same verb of each group's intentions. For example, if the participants would like to drink coffee without interrupting the previous intention of "drink milk", they would choose the cup for "Drink Coffee" due to the memory of the last intention's selection. "Take Medicine" as an intention can weaken this effect. Each participant is required to participate in two experiments for each intention.

In each experiment, five trials of five scene images with slightly different object distributions are set up (the different scene images are set up to prevent human memory processes). Up to 10 objects are included in the images. For each trial, the text of the intention name is displayed in the upper right corner of each image, i.e., trials are required for each subject. In each trial, participants are required to gaze at the intention stimulus in the upper right corner of the test scene for 1-2 s and then freely view the desired object. After finishing viewing, thumbs up is given, and the experimenter switches to the next scene image until all trials of a group of experiments are completed, recalibrates, and the next group of experiments is conducted. 282 intention-object pair data $\{I^i : O^1, O^2, \ldots, O^j\}$ (the participant views $O^1, O^2, \ldots, O^j$ for I^i) are collected. For each intention, the number of times that each object is viewed is counted, and the frequency of each object being viewed is taken as the conditional probability.

1) Priority Validation of Intention-Interpreting Object: Fig.4 (a) and (b) illustrate the results of the experiments for the pattern validation of object selection. We can observe that the probability of the cup is higher with "Drink a Cup of Milk" than with "Drink Milk," because the noun "cup" appears directly in the intention stimulus and serves as a reminder to the participant. The same pattern is also reflected in the second experiment (see Fig.4(c) (d)). Compared with "Drink Coffee,"



(a) Pupil Core eye tracker (b) Scene capturing and pupil detection (c) Gaze estimation

(d) Experiment Setup (e) Different test scene with intention of stimulation

Fig. 3. Experiment setup for pattern validation and gaze-based intention inference. (a) The used eye tracker consists of 1) an eye camera, 2) nose support, and 3) a scene camera. (b) The upper is the scene provided by the world camera, and the lower is pupil detection by the eye camera. (c) Gaze estimation through pupil image and scene image. (d) Experiment setup. (e) Several different test scenes. When performing pattern validation, the blue ellipse in the top right corner indicates the intended stimulus.

"Drink a cup of coffee" and "Drink a bowl of coffee" increase the probability of selecting the cup and bowl, respectively. This is consistent with human intuition, similar to the human conditioned reflex process, where an object corresponding to the nouns of the intention is a stimulus to remind people to choose it more frequently.

In this work, the objects corresponding to the nouns are viewed as intention-interpreting objects with the highest priority among intention-associated objects in intention expression. Although the existence of intention-interpreting objects is verified through experiments, it is highlighted that the intention-interpreting object for each intention does not require any data to be learned and is directly identified through human prior knowledge.

2) Non-Conditional Independence in Intention Inference: Fig.4 (c) and (d) show that people tend to focus on the object corresponding to the nouns, such as the cup or bowl. For example, as the cup (bowl) is corresponding to the nouns "cup (bowl)" of "Drink A Cup (Bowl) of Coffee," the frequency of the cup (bowl) increases for "Drink A Cup (Bowl) of Coffee" than "Drink Coffee." Besides, if people wish to drink a cup (bowl) of coffee, they tend to focus on the cup (bowl) and the coffee together. This indicates simultaneous selection. If people want to drink water, they tend to focus on the cup or the bowl because both the cup and bowl are liquid carriers, which play the same role in "Drink Water." This indicates the mutual exclusion of object selection; Both the simultaneous selection and the mutual exclusion selection indicate dependencies among features, which implies that the conditional independence assumption in the naïve Bayes may not be satisfied. Increasing the weight of predictive objects for specific intentions is necessary to weaken conditional independence.

TABLE II
ID NUMBER FOR 10 OBJECTS AND 10 INTENTIONS

ID	Objects	ID	Intentions
O1	Kettle	I1	Drink Water
O2	Cup	I2	Take Medicine
O3	Bowl	I3	Drink Milk
O4	Medicine	I4	Drink Coffee
O5	Yogurt	I5	Drink Coffee with Milk
O6	Milk	I6	Drink Yogurt
O7	Coffee	I7	Eat an Apple
O8	Banana	I8	Eat a Banana
O9	Apple	I9	Eat Fruit Salad
O10	Alarm	I10	Ask for Help

From the above, two observations can be obtained as follows:

- The objects corresponding to nouns, i.e., intention-interpreting objects, have the highest priority.
- There is simultaneous selection or mutual exclusion of subjects' attention objects at the time of intention expression. This also suggests that conditional independence in the naïve Bayes may not be satisfied.

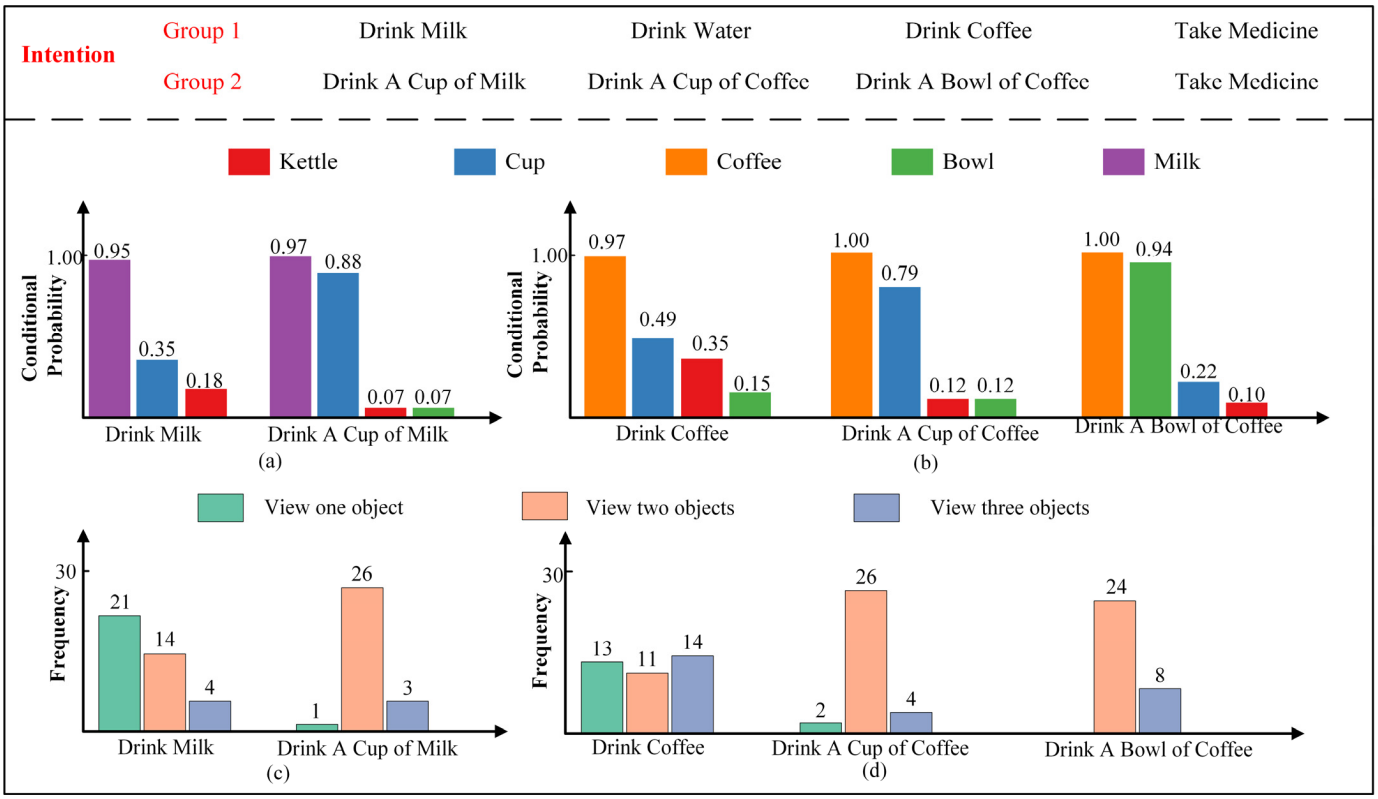


Fig. 4. Experimental results of pattern validation. (a) (b) Statistical results of object selection. The vertical axis is the conditional probability of the corresponding object being viewed. (c) (d) Experimental results of exploring the number of objects viewed for each intention. The horizontal axis indicates the different intentions, and the vertical axis indicates the viewing frequency of one or multiple objects marked in different colors.

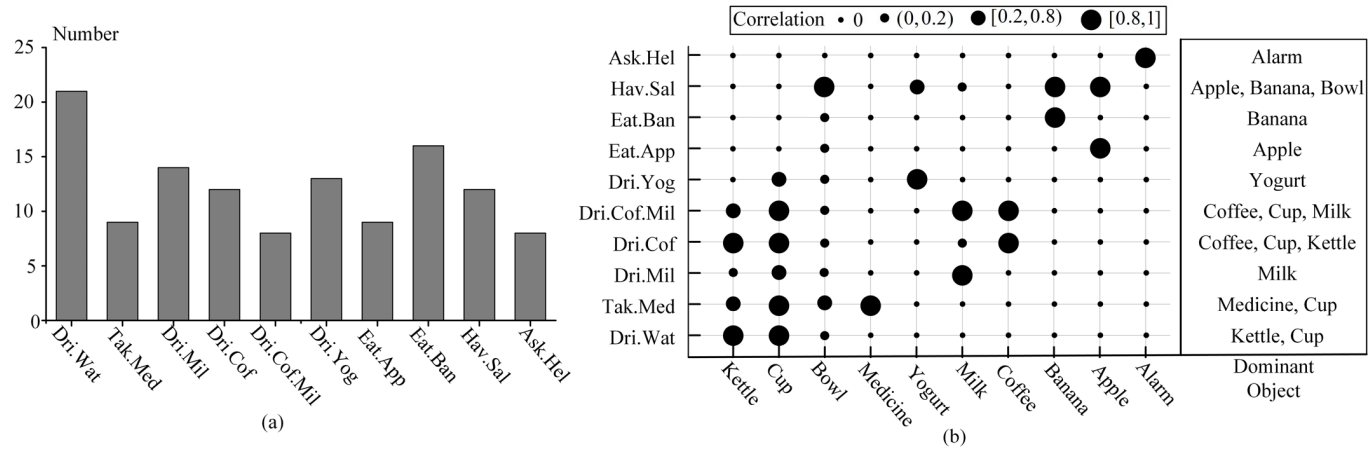


Fig. 5. Statistical results of building the intention database. (a) Distribution of intentions. The horizontal axis is the intention, and the vertical axis is the number of each intention. (b) Correlation plot between intentions and objects. The larger the dot, the stronger the correlation strength. The right side is the dominant object for each intention.

C. Construction of Intention Database

After finishing the pattern validation, the experiment on intention inference is carried out on our collected database, which contains 10 objects and 10 intentions, as shown in Table II. The intention representation in Table II follows the format "verb + noun", which is typically more general than the representation of "verb + quantifier + noun" in real scenarios [35]. 10 volunteers aged 20 to 25 with good eyesight are invited to participate in 10 trials for 10 intentions. For each trial, we first instruct the participants about the types of intention. Then the participants are asked to gaze at the intended objects sequentially based on their preferences. During each trial, the eye tracker can record the gaze sequence, which is used to analyze the intention-object pair $\{I^i : O^1, O^2, \ldots, O^j\}$.

To improve the efficiency of data collection, another 23 participants with short-sightedness are invited to participate in the survey. In the survey, the same objects and intentions are provided to each participant, and each participant selects objects to accomplish one type of intention. Both experimental collection and using the survey provide intention-object pair

TABLE III
WEIGHT MATRIX W IN CAWNB

W	O1	O2	O3	O4	O5	O6	O7	O8	O9	O10
I1	0.788	0.250	0.790	0.997	0.645	0.945	0.978	0.702	0.663	0.745
I2	0.888	0.965	0.049	0.932	0.544	0.693	0.905	0.755	0.748	0.762
I3	0.247	0.468	0.649	0.880	0.500	0.953	0.979	0.658	0.933	0.837
I4	0.865	0.393	0.407	0.997	0.856	0.931	0.960	0.264	0.382	0.418
I5	0.195	0.753	0.300	0.361	0.361	0.989	0.982	0.265	0.066	0.968
I6	0.518	0.183	0.810	0.973	0.973	0.858	0.640	0.823	0.957	0.854
I7	0.786	0.555	0.685	0.940	0.940	0.641	0.927	0.870	0.982	0.675
I8	0.665	0.801	0.652	0.923	0.926	0.848	0.734	0.946	0.989	0.681
I9	0.719	0.754	0.191	0.943	0.943	0.941	0.889	0.971	0.606	0.936
I10	0.484	0.604	0.985	0.615	0.615	0.450	0.456	0.542	0.734	0.699

data. Three hundred thirty intention-object pair data are collected by these two methods, which are mixed with computing the $P(O|I)$. What's more, to get $P(I)$, we collect the number of times that a group aged 20 to 30 execute the intentions in the database in 7 days.

Results: The distribution of intentions is shown in Fig.5(a). The frequency of each intention is taken as $P(I)$. Besides, we perform a statistical analysis of 330 intention-object pairs, and the conditional probability $P(O|I)$ is approximated by the frequency of selecting each object. A correlation plot (see Fig.5(b)) is used to present the database for the convenience of display and perception. The establishment of naïve Bayes, which is also the untrained class-specific attribute weighted naïve Bayes, is completed. Fig.5(b) shows that even though different participants view different objects, some objects are viewed by most of the participants. Here, the object that more than 80% of the participants selected for a specific intention is regarded as the dominant object of this intention.

D. Evaluation of the Proposed Intention Inference Approach

We detail the use of the proposed model for the collected data. First, the prior knowledge model of the intention-interpreting object is set up by prior knowledge. The intention-interpreting object matrix E is denoted as

$$\begin{array}{l} I_1 \rightarrow \\ \\ \\ \\ \\ - \\ \\ \\ \\ \\ I_{10} \rightarrow \\ \\ \end{array} \left\{ \begin{array}{cccccccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right\} = \mathrm{E}, \\ \begin{array}{cccccccccc} \uparrow & & & & & & & & & \uparrow \\ O_1 & & & & & - & & & & O_{10} \end{array} \quad (16)$$

where the column index of the matrix indicates the corresponding object and the row index indicates the corresponding intention.

What's more, 70% of the intention object pairs are used to train the weight matrix **W** in the CAWNB. The process of optimizing the objective function to convergence is shown in Fig.6. The well-trained weight matrix **W** is given in Table III.

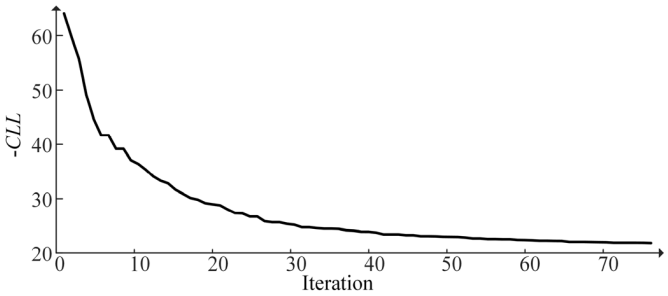


Fig. 6. *CLL* optimization convergence graph. The vertical axis is the value of the objective function *-CLL(w)*, and the horizontal axis is the number of iterations.

Some values in the matrix are equal because we have taken an approximation.

E. Comparison With Other Methods

To verify the effectiveness of the proposed method, we implement and evaluate the following baseline intention inference method:

1) Naïve Bayes [18]: The construction of the original database is the same as ours, and it is the base model of the proposed P-WNB. The difference from our method is that naïve Bayes only trains the conditional and intention prior probabilities. In contrast, the proposed method trains the conditional probabilities, the prior probabilities, and weights between intentions and objects. In addition, our method introduces the punished factor when computing posterior probabilities of intentions.

2) Support Vector Machine (SVM) [15]: The SVM classifier treats each object as a feature. Thus, the SVM classification results are 10 categories on 10 feature numbers in the proposed intention-object database. We implement a linear-kernel-based SVM classifier in our experiment to compare with our prosed method.

3) Decision Tree [19]: We implement the CART decision tree using the Gini coefficient by treating the objects contained in the knowledge base are viewed as decision features and the intention label as the final classification result.

4) Random Forest [19]: Random Forest is a machine learning algorithm that integrates a set of decision trees. We use Matlab embedded *TreeBagger* random forest function and set the number of trees $ntree = 20$.

5) Adaboost [19]: Adaboost integrates a set of weak classifiers to obtain a strong classifier. The Adaboost model is fitted using MATLAB function *fitensemble*. The number of weak classifiers is set to 100.

6) CFWNB [24]: CFWNB weights each attribute through the difference between feature-class (object-intention) correlation and the average feature-feature intercorrelation (object-object).

7) Hidden Naïve Bayes (HNB) [36]: HNB creates a hidden parent for each attribute, which considers the influence of all features.

Results: The results are reported by the mean accuracy and standard deviation of ten runs of a random 30% test

TABLE IV
COMPARISON WITH OTHER BASELINE METHODS

Method	Mean Accuracy (%)	STD (%)
Naïve Bayes [18]	93.82	2.14
SVM [15]	97.17	1.49
Decision Tree [19]	96.06	2.75
Random Forest [19]	97.58	1.36
Adaboost [19]	93.94	2.28
HNB [36]	95.45	2.25
CFWNB [24]	96.22	1.58
P-WNB	**98.43**	**1.09**

data, as shown in Table IV. The inference accuracies of the proposed P-WNB are 98.43 ± 1.09(mean accuracy (%) ± standard deviation (%)), which outperforms the state-of-the-art methods with higher mean accuracy and lower standard deviation. The results demonstrate that all methods perform well for "Ask for help" and "Take Medicine" due to their dissimilar dominant objects. However, the inference accuracy of the existing methods is low for "Eat Fruit," "Eat Fruit Salad," and other intentions with similar dominant objects. This is because the proposed method extracts the high-level features of each object rather than the simple semantic features.

F. Ablation Study

Ablative studies are conducted to analyze the contribution of the critical components of our proposed model. The detailed results of the ablation study and comparison with the improved naïve Bayes are shown in Table V. The results of CFWNB and HNB are given again for better comparison with improved naïve Bayes.

1) Benefit of Using Weighted Naïve Bayes: We evaluate the benefits of introducing CAWNB into intention inference models from experimental results. As shown in Table V, the mean inference accuracies of CFWNB and CAWNB are 96.22% and 96.85%, respectively. The good performance of CAWNB indicates that 1) the CAWNB assigning strategy, i.e., each object a specific weight to the different class, is significant; 2) the wrapper approach (used in CAWNB) has better classification performance than the filter approach (used in CFWNB). This is also consistent with intention inference models in which each object has a unique correlation strength to each intention. Additionally, the CAWNB model is learned iteratively from training data, making it more consistent with the actual data and compensating for the strong independence assumptions of the naïve Bayes.

2) Benefit of Using the Punished Factor: From the experimental results, it can be observed that adding the punished factor to the naïve Bayes improves the inference accuracy by nearly 2%. The advantage of the punished factor is that it takes into account the human conditioned reflex process, i.e., "When a specific intention is desired, objects associated with the name of the intention will most likely be attended to." By modeling the highest priority object as the punished factor, the accuracy of the inference is improved, and the algorithm is less sensitive to training data with errors.

By combining the advantages of the aforementioned two approaches, the proposed P-WNB, considering both data-driven by data training and model-driven by human prior knowledge, thereby achieves superior performance.

TABLE V
RESULTS OF ABLATION STUDY AND COMPARISON WITH IMPROVED NAÏVE BAYES

Method	Mean Accuracy (%)	STD (%)
Naïve Bayes [18]	93.82	2.14
HNB [36]	95.45	2.25
CFWNB [24]	96.22	1.58
Ours (NB. w. W)	96.85	1.28
Ours (NB. w. P)	95.39	1.63
Ours (NB. w. W & P)	**98.43**	**1.09**

NB: naïve Bayes.
NB. w. W: naïve Bayes with weights (CAWNB).
NB. w. P: naïve Bayes with punished factor.
NB. w. W & P: P-WNB

TABLE VI
COMPARISON RESULTS ON THE MIXED DATASET

Method	Mean Accuracy (%)	STD (%)
Naïve Bayes [18]	84.81	2.41
SVM [15]	86.32	1.59
Decision Tree [19]	85.57	2.13
Random Forest [19]	85.95	2.08
Adaboost [19]	85.46	2.06
CFWNB [24]	85.35	1.72
HNB [36]	85.62	1.98
P-WNB	**87.51**	**1.97**

G. Results on Challenging Task

In order to further verify the effectiveness of the proposed method, the data collected in Section IV-B and Section IV-C are integrated to obtain a mixed database, which poses a more challenging task than the case in Section IV-C.

We evaluate the proposed method on the mixed datasets and present the detailed comparison results with other methods as shown in Table VI. The average inference accuracy of the proposed P-WNB is 87.51 ± 1.97. Compared with the results in Table IV, the performance performed on mixed datasets decreases, which implies that ambiguous intentions yield challenges in intention inference. The best inference accuracy and relatively small standard deviation also verify the superiority of the proposed P-WNB method.

V. DISCUSSION

A. Advantages of the Proposed Method

Like several existing works, this study leverages eye-movement interaction to infer human intention from the patterns of human eye movement and scene information, without requiring excessive muscle movement. This approach provides novel ideas for the development of assistive devices.

The proposed approach differs from existing methods, which rely solely on data-driven machine-learning methods, by incorporating a prior knowledge model. Data-driven approaches have the apparent drawbacks of being highly reliant on big data and sensitive to noisy data, which yields low performance in real-life scenarios.

The prior knowledge model considers the process of human conditioned reflexes, i.e., when a person wants to express an intention, the objects corresponding to the nouns of the intention are equivalent to reminding the person to choose them. In this case, the objects corresponding to the nouns of the intention come into people's minds. For example, when a person wants to prepare a cup of coffee, coffee and cup remind the person to select them. Thus, the person is prone to select these two objects, even though the kettle is also necessary for this intention. This phenomenon of object preference can be interpreted as the different priorities given to different objects for each intention.

The object with the highest priority, called the intention-interpreting object, is served as the punished factor. The proposed method combines the punished factor with a superior performing class-specific attribute weighted naïve Bayes model. This approach models the priority of objects not reflected in the database, allowing intention inference to be more consistent with human common sense. Experimental results show that this approach achieves the best performance.

The problem of intention inference for the intentions with many common objects remains unresolved by prior works. As each type of intention has a unique set of intention-interpreting objects, the proposed approach provides a good solution to the problem of intention inference with many common objects. Although the dominant object somewhat embodies this idea, it does not model it well. Besides defining the dominant object by the high probability of being selected in the experiment, there will still be a duplicate part of the dominant object with different intentions.

In contrast, the intention-interpreting object can be considered the object with the highest priority among the dominant objects, which does not require any experimental data to learn. For different intentions, the more detailed the intention's name, the more comprehensive the setting of the intention-interpreting object in the prior knowledge model, and the clearer the discriminatory boundaries of different intentions will be. Introducing the prior knowledge model into intention inference improves the model's reasoning and discriminative power.

B. Limitations

Despite the positive results achieved in this study, the proposed P-WNB still has two major limitations. This study only considers whether one object is viewed or not, rather than the duration of viewing each object. It is desirable to take into account the duration of object-gazing and expand the range of model input to [0, 1], instead of being restricted to {0, 1}. Such processing may facilitate the generalization ability of the intention inference model. On the other hand, the types of intentions and objects involved in this study are relatively limited in comparison to those that occur in real life.

VI. Conclusion

In this paper, a hybrid-driven gaze-based intention inference method is proposed, which combines both the data-driven weighted naïve Bayes model and the model-driven prior knowledge model. The proposed approach leverages human prior knowledge to extract the intention-interpreting objects and introduces a data-driven weighted naïve Bayes model to represent the correspondence between intentions and objects. The experimental results demonstrate that our method achieves inference accuracies of 98.43% and 87.51% on two collected datasets respectively, which outperforms state-of-the-art methods. This approach a) tailors to disability applications and b) provides new ideas for designing assistive robots to help the disabled to accomplish their activities of daily living independently. In the future, a larger intention database, which contains more subjects and intentions in ample scenarios, will be constructed to test the proposed method.

References

[1] C. Guger, G. Edlinger, W. Harkam, I. Niedermayer, and G. Pfurtscheller, "How many people are able to operate an EEG-based brain–computer interface (BCI)?" *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 11, no. 2, pp. 145–147, Jun. 2003, doi: 10.1109/TNSRE.2003.814481.

[2] M. R. M. Tomari, Y. Kobayashi, and Y. Kuno, "Development of smart wheelchair system for a user with severe motor impairment," *Proc. Eng.*, vol. 41, pp. 538–546, Jan. 2012, doi: 10.1016/j.proeng.2012.07.209.

[3] D. Ao, R. Song, and J. Gao, "Movement performance of human–robot cooperation control based on EMG-driven hill-type and proportional models for an ankle power-assist exoskeleton robot," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 8, pp. 1125–1134, Aug. 2017, doi: 10.1109/TNSRE.2016.2583464.

[4] L. Chen, W. Su, Y. Feng, M. Wu, J. She, and K. Hirota, "Two-layer fuzzy multiple random forest for speech emotion recognition in human–robot interaction," *Inf. Sci.*, vol. 509, pp. 150–163, Jan. 2020, doi: 10.1016/j.ins.2019.09.005.

[5] C. Escolano, J. M. Antelis, and J. Minguez, "A telepresence mobile robot controlled with a noninvasive brain–computer interface," *IEEE Trans. Syst., Man, Cybern., B Cybern.*, vol. 42, no. 3, pp. 793–804, Jun. 2012, doi: 10.1109/TSMCB.2011.2177968.

[6] L. R. Hochberg et al., "Reach and grasp by people with tetraplegia using a neurally controlled robotic arm," *Nature*, vol. 485, no. 7398, pp. 372–375, May 2012, doi: 10.1038/nature11076.

[7] H. Park, S. Lee, M. Lee, M.-S. Chang, and H.-W. Kwak, "Using eye movement data to infer human behavioral intentions," *Comput. Hum. Behav.*, vol. 63, pp. 796–804, Oct. 2016, doi: 10.1016/j.chb.2016.06.016.

[8] B. Xu, J. Li, Y. Wong, Q. Zhao, and M. S. Kankanhalli, "Interact as you intend: Intention-driven human-object interaction detection," *IEEE Trans. Multimedia*, vol. 22, no. 6, pp. 1423–1432, Jun. 2020, doi: 10.1109/TMM.2019.2943753.

[9] X. Wang, X. Zheng, W. Chen, and F.-Y. Wang, "Visual human–computer interactions for intelligent vehicles and intelligent transportation systems: The state of the art and future directions," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 51, no. 1, pp. 253–265, Jan. 2021, doi: 10.1109/TSMC.2020.3040262.

[10] P. Wei, Y. Liu, T. Shu, N. Zheng, and S. C. Zhu, "Where and Why are They Looking? Jointly inferring human attention and intentions in complex tasks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 6801–6809, doi: 10.1109/CVPR.2018.00711.

[11] Z. Nan et al., "Learning to infer human attention in daily activities," *Pattern Recognit.*, vol. 103, Jul. 2020, Art. no. 107314, doi: 10.1016/j.patcog.2020.107314.

[12] Y. Fang et al., "Dual attention guided gaze target detection in the wild," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Feb. 2021, pp. 11385–11394, doi: 10.1109/CVPR46437.2021.01123.

[13] Z. Wan, C. Xiong, W. Chen, H. Zhang, and S. Wu, "Pupil-contour-based gaze estimation with real pupil axes for head-mounted eye tracking," *IEEE Trans. Ind. Informat.*, vol. 18, no. 6, pp. 3640–3650, Jun. 2022, doi: 10.1109/TII.2021.3118022.

[14] M. A. Eid, N. Giakoumidis, and A. El Saddik, "A novel eye-gaze-controlled wheelchair system for navigating unknown environments: Case study with a person with ALS," *IEEE Access*, vol. 4, pp. 558–573, 2016, doi: 10.1109/ACCESS.2016.2520093.

[15] F. Koochaki and L. Najafizadeh, "Predicting intention through eye gaze patterns," in *Proc. IEEE Biomed. Circuits Syst. Conf. (BioCAS)*, Oct. 2018, pp. 1–4, doi: 10.1109/BIOCAS.2018.8584665.

[16] F. Koochaki and L. Najafizadeh, "Eye gaze-based early intent prediction utilizing CNN-LSTM," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., (EMBS)*, Jul. 2019, pp. 1310–1313, doi: 10.1109/EMBC.2019.8857054.

[17] S. Li and X. Zhang, "Implicit human intention inference through gaze cues for people with limited motion ability," in *Proc. IEEE Int. Conf. Mechatronics Autom.*, Aug. 2014, pp. 257–262, doi: 10.1109/ICMA.2014.6885705.

[18] S. Li and X. Zhang, "Implicit intention communication in human–robot interaction through visual behavior studies," *IEEE Trans. Hum.-Mach. Syst.*, vol. 47, no. 4, pp. 437–448, Aug. 2017, doi: 10.1109/THMS.2017.2647882.

[19] F. Koochaki and L. Najafizadeh, "A data-driven framework for intention prediction via eye movement with applications to assistive systems," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 974–984, 2021, doi: 10.1109/TNSRE.2021.3083815.

[20] C.-M. Huang, S. Andrist, A. Sauppé, and B. Mutlu, "Using gaze patterns to predict task intent in collaboration," *Frontiers Psychol.*, vol. 6, p.1049, Jul. 2015, doi: 10.3389/fpsyg.2015.01049.

[21] H. Zhang, L. Jiang, and L. Yu, "Class-specific attribute value weighting for naive Bayes," *Inf. Sci.*, vol. 508, pp. 260–274, Jan. 2020, doi: 10.1016/j.ins.2019.08.071.

[22] L. Jiang, C. Li, S. Wang, and L. Zhang, "Deep feature weighting for naive Bayes and its application to text classification," *Eng. Appl. Artif. Intell.*, vol. 52, pp. 26–39, Jun. 2016, doi: 10.1016/j.engappai.2016.02.002.

[23] M. Hall, "A decision tree-based attribute weighting filter for naive Bayes," *Knowledge-Based Syst.*, vol. 20, no. 2, pp. 120–126, Mar. 2007, doi: 10.1016/j.knosys.2006.11.008.

[24] L. Jiang, L. Zhang, C. Li, and J. Wu, "A correlation-based feature weighting filter for naive Bayes," *IEEE Trans. Knowl. Data Eng.*, vol. 31, no. 2, pp. 201–213, Feb. 2019, doi: 10.1109/TKDE.2018.2836440.

[25] H. Zhang and S. Sheng, "Learning weighted naive Bayes with accurate ranking," in *Proc. 4th IEEE Int. Conf. Data Mining (ICDM)*, Nov. 2004, pp. 567–570, doi: 10.1109/ICDM.2004.10030.

[26] C. H. Lee, F. Gutierrez, and D. Dou, "Calculating feature weights in naive Bayes with Kullback–Leibler measure," in *Proc. IEEE 11th Int. Conf. Data Mining*, Dec. 2011, pp. 1146–1151, doi: 10.1109/ICDM.2011.29.

[27] J. Wu, Z. Cai, X. Chen, S. Ao, and Y. Zhang, "Attribute weighting via differential evolution for attribute weighted clonal selection algorithm," *J. Comput.*, vol. 7, no. 12, pp. 1–7, Dec. 2012.

[28] L. Jiang, L. Zhang, L. Yu, and D. Wang, "Class-specific attribute weighted naive Bayes," *Pattern Recognit.*, vol. 88, pp. 321–330, Apr. 2019, doi: 10.1016/j.patcog.2018.11.032.

[29] L. K. Foo, S. L. Chua, and N. Ibrahim, "Attribute weighted naïve Bayes classifier," *Comput., Mater. Continua*, vol. 71, no. 1, pp. 1945–1957, 2022, doi: 10.32604/cmc.2022.022011.

[30] E. Frank, M. Hall, and B. Pfahringer, "Locally weighted naive Bayes," 2012, *arXiv:1212.2487*.

[31] C. Zhu, R. H. Byrd, P. Lu, and J. Nocedal, "Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization," *ACM Trans. Math. Softw.*, vol. 23, no. 4, pp. 550–560, Dec. 1997, doi: 10.1145/279232.279236.

[32] J. Zhang and A. C. Sanderson, "JADE: Adaptive differential evolution with optional external archive," *IEEE Trans. Evol. Comput.*, vol. 13, no. 5, pp. 945–958, Oct. 2009, doi: 10.1109/TEVC.2009.2014613.

[33] T. Chen, R. Chen, L. Nie, X. Luo, X. Liu, and L. Lin, "Neural task planning with AND–OR graph representations," *IEEE Trans. Multimedia*, vol. 21, no. 4, pp. 1022–1034, Apr. 2019, doi: 10.1109/TMM.2018.2870062.

[34] M. Kassner, W. Patera, and A. Bulling, "Pupil: An open source platform for pervasive eye tracking and mobile gaze-based interaction," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput., Adjunct Publication*, 2014, pp. 1151–1160, doi: 10.1145/2638728.2641695.

[35] G. A. Sigurdsson, O. Russakovsky, and A. Gupta, "What actions are needed for understanding human actions in videos?" in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2137–2146, doi: 10.1109/ICCV.2017.235.

[36] L. Jiang, H. Zhang, and Z. Cai, "A novel Bayes model: Hidden naive Bayes," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 10, pp. 1361–1371, Oct. 2009, doi: 10.1109/TKDE.2008.234.