

CHATBOT CONFIGURATION METHOD AND APPARATUS, COMPUTER DEVICE, AND STORAGE MEDIUM

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims priority to Chinese Patent Application No. 201910676824.3, filed with the National Intellectual Property Administration, PRC on July 25, 2019 and entitled "CHATBOT CONFIGURATION METHOD AND APPARATUS, COMPUTER DEVICE, AND STORAGE MEDIUM", which is incorporated herein by reference in its entirety.

TECHNICAL FIELD

This application relates to a chatbot configuration method and apparatus, a computer device, and a storage medium.

BACKGROUND

With the development of computer technologies, an optical character recognition (OCR)-based service application method emerges. OCR refers to a process in which an electronic device (for example, a scanner or a digital camera) examines printed characters on paper, determines shapes of the characters by detecting dark and light areas, and translates the shapes into computer text by using a character recognition method. The OCR-based service application method includes capturing, through OCR, content from a paper form previously filled in by a user and automatically entering the content in a system.

However, the inventor realizes that the current OCR-based service application method still requires a user to fill in a paper form first, and automates only a subsequent entry process. Consequently, there is a problem of low service processing efficiency.

SUMMARY

According to various embodiments disclosed in this application, a chatbot configuration method and apparatus, a computer device, and a storage medium are provided.

A chatbot configuration method includes:

obtaining a service form scanning image, extracting form feature information in the service form scanning image, determining a field type of a literal block in the form feature information, identifying, based on the field type, a target literal block whose field needs to be filled in, and establishing an association relationship between literal blocks based on the target literal block whose field needs to be filled in.

determining, by querying a preset data type configuration table based on the target literal block whose field needs to be filled in, a data type of a field that needs to be filled in and that corresponds to each target literal block;

generating a service segment statement of each target literal block based on the data type of the field that needs to be filled in and that corresponds to each target literal block, the association relationship between literal blocks, and a preset statement template; and

configuring a chatbot based on location information of each target literal block in the form feature information and the service segment statement of each target literal block.

A chatbot configuration apparatus includes:

an obtaining module, configured to obtain a service form scanning image, extract form feature information in the service form scanning image, determine a field type of a literal block in the form feature information, identify, based on the field type, a target literal block whose field needs to be filled in, and establish an association relationship between literal blocks based on the target literal block whose field needs to be filled in;

a first processing module, configured to determine, by querying a preset data type configuration table based on the target literal block whose field needs to be filled in, a data type of a field that needs to be filled in and that corresponds to each target literal block;

a second processing module, configured to generate a service segment statement of each target literal block based on the data type of the field that needs to be filled in and that corresponds to each target literal block, the association relationship between literal blocks, and a preset statement template; and

a configuration module, configured to configure a chatbot based on location information of each target literal block in the form feature information and the service segment statement of each

target literal block.

A computer device includes a memory and one or more processors, where the memory stores a computer-readable instruction, and when the computer-readable instruction is executed by the one or more processors, the one or more processors are enabled to perform the following steps:

obtaining a service form scanning image, extracting form feature information in the service form scanning image, determining a field type of a literal block in the form feature information, identifying, based on the field type, a target literal block whose field needs to be filled in, and establishing an association relationship between literal blocks based on the target literal block whose field needs to be filled in.

determining, by querying a preset data type configuration table based on the target literal block whose field needs to be filled in, a data type of a field that needs to be filled in and that corresponds to each target literal block;

generating a service segment statement of each target literal block based on the data type of the field that needs to be filled in and that corresponds to each target literal block, the association relationship between literal blocks, and a preset statement template; and

configuring a chatbot based on location information of each target literal block in the form feature information and the service segment statement of each target literal block.

One or more nonvolatile computer-readable storage media that store a computer-readable instruction are provided, where when the computer-readable instruction is executed by one or more processors, the one or more processors are enabled to perform the following steps:

obtaining a service form scanning image, extracting form feature information in the service form scanning image, determining a field type of a literal block in the form feature information, identifying, based on the field type, a target literal block whose field needs to be filled in, and establishing an association relationship between literal blocks based on the target literal block whose field needs to be filled in.

determining, by querying a preset data type configuration table based on the target literal block whose field needs to be filled in, a data type of a field that needs to be filled in and that corresponds to each target literal block;

generating a service segment statement of each target literal block based on the data type of the field that needs to be filled in and that corresponds to each target literal block, the association relationship between literal blocks, and a preset statement template; and

configuring a chatbot based on location information of each target literal block in the form feature information and the service segment statement of each target literal block.

The details of the one or more embodiments of this application are set forth in the following accompanying drawings and description. Other features and advantages of this application become clear from the specification, the accompanying drawings, and the claims.

BRIEF DESCRIPTION OF DRAWINGS

To describe the technical solutions in the embodiments of this application more clearly, the following briefly describes the accompanying drawings required for describing the embodiments. Clearly, the accompanying drawings in the following description are merely some embodiments of this application, and a person of ordinary skill in the art can further derive other drawings from these accompanying drawings without creative efforts.

FIG. 1 is a diagram of an application scenario of a chatbot configuration method according to one or more embodiments;

FIG. 2 is a schematic flowchart of a chatbot configuration method according to one or more embodiments;

FIG. 3 is a schematic diagram of a subflow of step 202 in FIG. 1 according to one or more embodiments;

FIG. 4 is a schematic diagram of a subflow of step 202 in FIG. 1 according to one or more embodiments;

FIG. 5 is a schematic diagram of a subflow of step 202 in FIG. 1 according to one or more embodiments;

FIG. 6 is a schematic diagram of a subflow of step 206 in FIG. 1 according to one or more embodiments;

FIG. 7 is a schematic diagram of a subflow of step 608 in FIG. 1 according to one or more embodiments;

FIG. 8 is a block diagram of a chatbot configuration apparatus according to one or more embodiments; and

FIG. 9 is a block diagram of a computer device according to one or more embodiments.

DESCRIPTION OF EMBODIMENTS

To make the technical solutions and advantages of this application more clearly, the following further describes this application in detail with reference to the accompanying drawings and embodiments. It should be understood that the specific embodiments described herein are merely intended to explain this application, and are not intended to limit this application.

A chatbot configuration method provided in this application may be applied to an application environment shown in FIG. 1. A terminal 102 communicates with a server 104 via a network. The server 104 obtains a service form scanning image from the terminal 102, extracts form feature information in the service form scanning image, determines a field type of a literal block in the form feature information, identifies, based on the field type, a target literal block whose field needs to be filled in, and establishes an association relationship between literal blocks based on the target literal block whose field needs to be filled in; determines, by querying a preset data type configuration table based on the target literal block whose field needs to be filled in, a data type of a field that needs to be filled in and that corresponds to each target literal block; generates a service segment statement of each target literal block based on the data type of the field that needs to be filled in and that corresponds to each target literal block, the association relationship between literal blocks, and a preset statement template; and configures a chatbot based on location information of each target literal block in the form feature information and the service segment statement of each target literal block. The terminal 102 may be but is not limited to various personal computers, laptops, smartphones, tablets, and portable wearable devices. The server 104 may be implemented by an independent server or a server cluster that includes multiple servers.

In an embodiment, as shown in FIG. 2, a chatbot configuration method is provided. Description is provided by using an example in which the method is applied to the server in FIG. 1. The method includes the following steps.

Step S202: Obtain a service form scanning image, extract form feature information in the

service form scanning image, determine a field type of a literal block in the form feature information, identify, based on the field type, a target literal block whose field needs to be filled in, and establish an association relationship between literal blocks based on the target literal block whose field needs to be filled in.

The service form scanning image may be a scanning copy of a paper service application form or a picture of an electronic service application form. The form feature information in the service form scanning image may be extracted by using a character recognition algorithm. In the character recognition algorithm, the form feature information in the service form scanning image may be extracted through feature extraction, text locating, and optical recognition. Feature extraction means using the service form scanning image as an input based on a trained target detection model in the character recognition algorithm and extracting a feature by using a convolutional neural network in the target detection model. Text locating means obtaining location information of each detected literal block picture and location information of each detected literal symbol picture based on features extracted by using the trained target detection model in the character recognition algorithm. Optical recognition means recognizing text and symbols in each detected literal block picture and each detected literal symbol picture by using a trained picture classification model in the character recognition algorithm.

The trained target detection model and the trained picture classification model in the character recognition algorithm are obtained through training based on a large quantity of sample pictures that include text and literal symbols. Text captured by the character recognition algorithm is in a form of block. For example, a field "Full Name" is captured as a block. The form feature information includes each literal block, the location information of each literal block, each literal symbol, and the location information of each literal symbol. The location information refers to coordinates of each literal block or each literal symbol relative to the entire picture, and is expressed in pixels. Coordinates of the top left corner and the bottom right corner of a capture box are (top, left) and (bottom, right), respectively. A location and a size of the capture box are determined based on the coordinates of the two points.

A field type of each literal block includes a field that needs to be filled in, an option field, and a

comment field. The option field and the comment field correspond to the field that needs to be filled in. The server determines, based on the field type of each literal block, that a literal block whose field type is a field that needs to be filled in is the target literal block, that a literal block whose field type is an option field is an option field literal block, and that a literal block whose field type is a comment field is a comment field literal block, then determines a correspondence among each target literal block, each option field literal block, and each comment field literal block based on a distance between each target literal block and each option field literal block, and a distance between each target literal block and each comment field literal block, and establishes the association relationship between literal blocks based on the correspondence.

Step S204: Determine, by querying a preset data type configuration table based on the target literal block whose field needs to be filled in, a data type of a field that needs to be filled in and that corresponds to each target literal block.

The server can determine, by querying the preset data type configuration table based on the target literal block whose field needs to be filled in, the data type of the field that needs to be filled in and that corresponds to each target literal block. In the data type configuration table, the data type of the field that needs to be filled in and that corresponds to each target literal block is preset. For example, when the field that needs to be filled in indicates a telephone number or age, the corresponding data type is number.

Step S206: Generate a service segment statement of each target literal block based on the data type of the field that needs to be filled in and that corresponds to each target literal block, the association relationship between literal blocks, and a preset statement template.

The server determines an option field literal block and a comment field literal block corresponding to each target literal block based on the association relationship between literal blocks, and generates the service segment statement of each target literal block based on the data type of the field that needs to be filled in and that corresponds to each target literal block, the preset statement template, and the option field literal block and the comment field literal block corresponding to each target literal block. The service segment statement refers to a sample dialogue segment for obtaining the field that needs to be filled in and that corresponds to each target

literal block, and includes a machine reply statement, a comment prompt statement, and a customer intention statement.

The machine reply statement is based on each target literal block. The machine reply statement refers to a term used by a chatbot to ask a customer about required information corresponding to the field that needs to be filled in. The comment prompt statement is based on the association relationship between literal blocks. When there is a comment field literal block corresponding to the target literal block, the comment prompt statement can be obtained based on the comment field literal block. The comment prompt statement is used to prompt the customer to input the required information corresponding to the field that needs to be filled in. The customer intention statement is based on the data type of each target literal block and the association relationship between literal blocks, and is a statement in which the customer provides the required information corresponding to the field that needs to be filled in. For example, when the data type is number, the obtained customer intention statement should be a string of numbers. The server obtains the service segment statement of each target literal block in a sequence of the machine reply statement, the comment prompt statement, and the user intention statement.

Step S208: Configure a chatbot based on location information of each target literal block in the form feature information and the service segment statement of each target literal block.

A filling sequence of fields that need to be filled in and that correspond to all the target literal blocks can be determined based on the location information of all the target literal blocks, and service segment statements of all the target literal blocks are stored based on the filling sequence of all the target literal blocks, to generate scenario process information of a complete service application, and the chatbot is configured based on the scenario process information.

According to the chatbot configuration method, the form feature information in the service form scanning image is extracted, the field type of the literal block in the form feature information is determined, the target literal block whose field needs to be filled in is identified based on the field type, and the association relationship between literal blocks is established based on the target literal block whose field needs to be filled in; the data type of the field that needs to be filled in and that corresponds to each target literal block is determined by querying the preset data type configuration

table; the service segment statement of each target literal block is generated based on the data type of the field that needs to be filled in and that corresponds to each target literal block, the association relationship between literal blocks, and the preset statement template; and the chatbot is configured based on the location information of each target literal block in the form feature information and the service segment statement of each target literal block. Therefore, service processing can be performed based on the configured chatbot, so that a user can provide required information in an original paper form through online chat to complete the service application, which improves service processing efficiency.

In an embodiment, as shown in FIG. 3, step S202 includes the following steps:

Step S302: Obtain the service form scanning image, and preprocess the service form scanning image.

Step S304: Obtain location information of each literal block and location information of each literal symbol in the preprocessed service form scanning image based on a trained target detection model, where the target detection model is obtained through training by using sample pictures that include literal blocks and literal symbols.

Step S306: Divide the service form scanning image based on the location information of each literal block and the location information of each literal symbol, to obtain multiple literal block images and multiple literal symbol images.

Step S308: Extract a literal block and a literal symbol in each literal block image and each literal symbol image based on a trained picture classification model, to obtain the form feature information in the service form scanning image, where the picture classification model is obtained through training by using sample pictures that include literal blocks and literal symbols.

Preprocessing includes denoising and slant correction. The target detection model is obtained through training by using the sample pictures that include the literal blocks and the literal symbols. After the service form scanning image is input into the trained target detection model in the character recognition algorithm, the convolutional neural network in the trained target detection model extracts features of the service form scanning image, and the location information of each literal block and the location information of each literal symbol in the service form scanning image

are obtained based on the extracted features and a fully connected layer in the trained target detection model; the service form scanning image is divided based on the location information of each literal block and the location information of each literal symbol, to obtain the multiple literal block images and the multiple literal symbol images; and finally, text in the images can be recognized by using the trained picture classification model. Both the target detection model and the picture classification model are obtained through training by using the sample pictures that include the literal blocks and the literal symbols. The target detection model may be common YOLO, Faster R-CNN, SSD, or the like. The picture classification model may be ResNet. Common literal symbols include a long underscore, a check box, and the like. These literal symbols can be used to help classify literal blocks.

In the foregoing embodiment, the preprocessed service form scanning image is processed by using the trained target detection model and the trained picture classification model, to extract the form feature information in the service form scanning image.

In an embodiment, step 304 includes:

inputting the preprocessed service form scanning image into the trained target detection model; extracting the features of the preprocessed service form scanning image based on the convolutional neural network in the target detection model; and

obtaining the location information of each literal block and the location information of each literal symbol in the preprocessed service form scanning image based on the fully connected layer in the target detection model and the features of the service form scanning image.

In an embodiment, as shown in FIG. 4, step S202 includes the following steps:

Step S402: Input each literal block in the form feature information into a trained classification model, to obtain confidence levels indicating that the literal blocks belong to preset field types.

Step S404: Determine a coordinate distance between each literal block and each literal symbol based on location information of each literal block and location information of each literal symbol in the form feature information.

Step S406: Use each literal symbol whose coordinate distance to each literal block falls within a preset distance threshold range as a literal symbol associated with the literal block.

Step S408: Revise the confidence levels indicating that the literal blocks belong to the preset field types based on an association between each literal block and each literal symbol.

Step S410: Sort the confidence levels indicating that the literal blocks belong to the preset field types, and use a field type with the highest confidence level as a field type of each literal block.

The server inputs each literal block in the form feature information into the trained classification model, to obtain the confidence levels indicating that the literal blocks belong to the preset field types. The confidence levels indicating that the literal blocks belong to the preset field types are used to indicate probabilities that each literal block belongs to the preset field types. After obtaining the confidence levels indicating that the literal blocks belong to the preset field types, the server determines the coordinate distance between each literal block and each literal symbol based on the location information of each literal block and the location information of each literal symbol in the form feature information; uses each literal symbol whose coordinate distance to each literal block falls within the preset distance threshold range as the literal symbol associated with the literal block; revises the confidence levels indicating that the literal blocks belong to the preset field types based on the association between each literal block and each literal symbol; and finally, sorts the confidence levels indicating that the literal blocks belong to the preset field types, and uses the field type with the highest confidence level as the field type of each literal block.

The revising the confidence levels indicating that the literal blocks belong to the preset field types based on the association between each literal block and each literal symbol means adjusting a confidence level of a literal block based on a type of an associated literal symbol when the literal block is associated with the literal symbol. For example, if a field follows a check box, a confidence level indicating that the field is "an option field" is increased; or if a field is followed by a long underscore, a confidence level indicating that the field is "a field that needs to be filled in" is increased. Further, a field that needs to be filled in includes a mandatory field and an optional field, and the field that needs to be filled in can be further classified by detecting whether there is a required symbol before or after a literal block.

In the foregoing embodiment, the confidence levels indicating that the literal blocks belong to the preset field types are obtained based on the trained classification model; the confidence levels

indicating that the literal blocks belong to the preset field types are revised based on the association between each literal block and each literal symbol; and the field type with the highest confidence level is used as the field type of each literal block. As such, the field type of each literal block is determined.

In an embodiment, as shown in FIG. 5, the field type includes a field that needs to be filled in, an option field, and a comment field, and step S202 includes the following steps:

Step S502: Determine, based on a field type of each literal block, that a literal block whose field type is a field that needs to be filled in is the target literal block whose field needs to be filled in.

Step S504: Determine a distance between each target literal block and each option field literal block, and a distance between each target literal block and each comment field literal block based on location information of each literal block in the form feature information.

Step S506: Determine an option field literal block and a comment field literal block corresponding to each target literal block based on the distance between each target literal block and each option field literal block, and the distance between each target literal block and each comment field literal block.

Step S508: Establish an association relationship between each target literal block and both the corresponding option field literal block and the corresponding comment field literal block.

The field type of each literal block includes the field that needs to be filled in, the option field, and the comment field. The option field and the comment field correspond to the field that needs to be filled in. The server determines, based on the field type of each literal block, that a literal block whose field type is a field that needs to be filled in is the target literal block whose field needs to be filled in, that a literal block whose field type is an option field is an option field literal block, and that a literal block whose field type is a comment field is a comment field literal block; determines the distance between each target literal block and each option field literal block, and the distance between each target literal block and each comment field literal block based on the location information of each literal block in the form feature information; determines the option field literal block and the comment field literal block corresponding to each target literal block based on the

distance between each target literal block and each option field literal block, and the distance between each target literal block and each comment field literal block; and establishes the association relationship between each target literal block and both the corresponding option field literal block and the corresponding comment field literal block.

In the foregoing embodiment, it is determined, based on the field type of each literal block, that the literal block whose field type is a field that needs to be filled in is the target literal block; the option field literal block and the comment field literal block corresponding to each target literal block are determined based on the location information of each literal block in the form feature information; and the association relationship between literal blocks is established. As such, the association relationship between literal blocks is determined.

In an embodiment, as shown in FIG. 6, step S206 includes the following steps:

Step S602: Determine, based on the location information of each target literal block in the form feature information, a filling sequence of fields that need to be filled in and that correspond to all the target literal blocks.

Step S604: Determine a service process sequence of service segment statements of all the target literal blocks based on the filling sequence of the fields that need to be filled in and that correspond to all the target literal blocks.

Step S606: Generate scenario process information of a service application based on the service process sequence.

Step S608: Configure the chatbot by performing model training based on the scenario process information and the service segment statements of all the target literal blocks.

The service process sequence can be obtained by integrating the service segment statements of all the target literal blocks based on the filling sequence of the fields that need to be filled in and that correspond to all the target literal blocks; the scenario process information of the service application can be generated based on the service process sequence and the service segment statements of all the target literal blocks; model training is performed based on the scenario process information and the service segment statements of all the target literal blocks, to obtain a natural language understanding model and a dialogue management model; and the chatbot is configured

based on the natural language understanding model and dialogue management model. The natural language understanding model is used to determine a user intention based on a user statement and capture entity information, and the dialogue management model is used to determine a reply statement based on the user statement and the user intention.

In the foregoing embodiment, the filling sequence of the fields that need to be filled in and that correspond to all the target literal blocks is determined based on the location information of each target literal block; the service process sequence of the service segment statements of all the target literal blocks is determined based on the filling sequence of the fields that need to be filled in and that correspond to all the target literal blocks; the scenario process information of the service application is generated; and the chatbot is configured by performing model training based on the scenario process information and the service segment statements of all the target literal blocks.

In an embodiment, as shown in FIG. 7, step S608 includes the following steps:

Step S702: Perform model training by inputting the service segment statements of all the target literal blocks into an initial natural language understanding model as a first training set, to obtain the natural language understanding model, where the natural language understanding model is used to determine the user intention based on the user statement and capture the entity information.

Step S704: Perform model training by inputting the scenario process information into an initial dialogue management model as a second training set, to obtain the dialogue management model, where the dialogue management model is used to determine the reply statement based on the user statement and the user intention.

Step S706: Configure the chatbot based on the natural language understanding model and the dialogue management model.

The service segment statements of all the target literal blocks include a machine reply statement, a comment prompt statement, and a user intention statement. Model training is performed by inputting the service segment statements of all the target literal blocks into the initial natural language understanding model as the first training set, so that the natural language understanding model can determine the user intention based on the user intention statement and capture required information in the user intention statement as the entity information. Model

training is performed by inputting the scenario process information into the initial dialogue management model as the second training set, so that the dialogue management model can determine the corresponding machine reply statement and the corresponding comment prompt statement based on the user statement and the user intention. The chatbot can be configured based on the natural language understanding model and the dialogue management model. After the chatbot is configured, in a task of the chatbot, when a customer enters a first user intention statement, the natural language understanding model determines a user intention based on the first user intention statement, and inputs the user intention into the dialogue management model. The dialogue management model determines a corresponding machine reply statement and a corresponding comment prompt statement based on the user intention, and pushes the machine reply statement and the comment prompt statement. The customer returns a second user intention statement based on the pushed machine reply statement and comment prompt statement. The natural language understanding model captures required information in the second user intention statement as entity information.

In the foregoing embodiment, the natural language understanding model is obtained based on the service segment statements of all the target literal blocks; the dialogue management model is obtained based on the scenario process information; and the chatbot is configured based on the natural language understanding model and the dialogue management model.

It should be understood that although the steps in the flowcharts in FIG. 2 to FIG. 7 are shown in a sequence indicated by arrows, these steps are not necessarily performed in the sequence indicated by the arrows. Unless expressly stated in this specification, these steps are not performed in a strict sequence, and these steps can be performed in other sequences. In addition, some of the steps in FIG. 2 to FIG. 7 may include multiple sub-steps or stages. These sub-steps or phases are not necessarily performed at the same time, but may be performed at different time. These sub-steps or stages are not necessarily sequentially performed, but may be alternately performed with at least some of other steps or sub-steps or stages of other steps.

In an embodiment, as shown in FIG. 8, a chatbot configuration apparatus is provided, and includes an obtaining module 802, a first processing module 804, a second processing module 806,

and a configuration module 808.

The obtaining module 802 is configured to obtain a service form scanning image, extract form feature information in the service form scanning image, determine a field type of a literal block in the form feature information, identify, based on the field type, a target literal block whose field needs to be filled in, and establish an association relationship between literal blocks based on the target literal block whose field needs to be filled in.

The first processing module 804 is configured to determine, by querying a preset data type configuration table based on the target literal block whose field needs to be filled in, a data type of a field that needs to be filled in and that corresponds to each target literal block.

The second processing module 806 is configured to generate a service segment statement of each target literal block based on the data type of the field that needs to be filled in and that corresponds to each target literal block, the association relationship between literal blocks, and a preset statement template.

The configuration module 808 is configured to configure a chatbot based on location information of each target literal block in the form feature information and the service segment statement of each target literal block.

The chatbot configuration apparatus extracts the form feature information in the service form scanning image, determines the field type of the literal block in the form feature information, identifies, based on the field type, the target literal block whose field needs to be filled in, and establishes the association relationship between literal blocks based on the target literal block whose field needs to be filled in; determines, by querying the preset data type configuration table, the data type of the field that needs to be filled in and that corresponds to each target literal block; generates the service segment statement of each target literal block based on the data type of the field that needs to be filled in and that corresponds to each target literal block, the association relationship between literal blocks, and the preset statement template; and configures the chatbot based on the location information of each target literal block in the form feature information and the service segment statement of each target literal block. Therefore, service processing can be performed based on the configured chatbot, so that a user can provide required information in an original paper

form through online chat to complete the service application, which improves service processing efficiency.

In an embodiment, the obtaining module is further configured to obtain the service form scanning image, and preprocess the service form scanning image; obtain location information of each literal block and location information of each literal symbol in the preprocessed service form scanning image based on a trained target detection model, where the target detection model is obtained through training by using sample pictures that include literal blocks and literal symbols; divide the service form scanning image based on the location information of each literal block and the location information of each literal symbol, to obtain multiple literal block images and multiple literal symbol images; and extract a literal block and a literal symbol in each literal block image and each literal symbol image based on a trained picture classification model, to obtain the form feature information in the service form scanning image, where the picture classification model is obtained through training by using sample pictures that include literal blocks and literal symbols.

In an embodiment, the obtaining module is further configured to input the preprocessed service form scanning image into the trained target detection model; extract features of the preprocessed service form scanning image based on a convolutional neural network in the target detection model; and obtain the location information of each literal block and the location information of each literal symbol in the preprocessed service form scanning image based on a fully connected layer in the target detection model and the features of the service form scanning image.

In an embodiment, the obtaining module is further configured to input each literal block in the form feature information into a trained classification model, to obtain confidence levels indicating that the literal blocks belong to preset field types; determine a coordinate distance between each literal block and each literal symbol based on location information of each literal block and location information of each literal symbol in the form feature information; use each literal symbol whose coordinate distance to each literal block falls within a preset distance threshold range as a literal symbol associated with the literal block; revise the confidence levels indicating that the literal blocks belong to the preset field types based on an association between each literal block and each literal symbol; and sort the confidence levels indicating that the literal blocks belong to the preset

field types, and use a field type with the highest confidence level as a field type of each literal block.

In an embodiment, the field type includes a field that needs to be filled in, an option field, and a comment field, and the obtaining module is further configured to determine, based on a field type of each literal block, that a literal block whose field type is a field that needs to be filled in is the target literal block whose field needs to be filled in; determine a distance between each target literal block and each option field literal block, and a distance between each target literal block and each comment field literal block based on location information of each literal block in the form feature information; determine an option field literal block and a comment field literal block corresponding to each target literal block based on the distance between each target literal block and each option field literal block, and the distance between each target literal block and each comment field literal block; and establish an association relationship between each target literal block and both the corresponding option field literal block and the corresponding comment field literal block.

In an embodiment, the configuration module is further configured to determine, based on the location information of each target literal block in the form feature information, a filling sequence of fields that need to be filled in and that correspond to all the target literal blocks; determine a service process sequence of service segment statements of all the target literal blocks based on the filling sequence of the fields that need to be filled in and that correspond to all the target literal blocks; generate scenario process information of a service application based on the service process sequence; and configure the chatbot by performing model training based on the scenario process information and the service segment statements of all the target literal blocks.

In an embodiment, the configuration module is further configured to perform model training by inputting the service segment statements of all the target literal blocks into an initial natural language understanding model as a first training set, to obtain a natural language understanding model, where the natural language understanding model is used to determine a user intention based on a user statement and capture entity information; perform model training by inputting the scenario process information into an initial dialogue management model as a second training set, to obtain a dialogue management model, where the dialogue management model is used to determine a reply

statement based on the user statement and the user intention; and configure the chatbot based on the natural language understanding model and the dialogue management model.

For specific limitations of the chatbot configuration apparatus, refer to the foregoing limitations of the chatbot configuration method. Details are not described herein again. All or some of the modules in the chatbot configuration apparatus may be implemented by software, hardware, or a combination thereof. The foregoing modules may be built in or independent of a processor in a computer device in a hardware form, or may be stored in a memory in a computer device in a software form, so that the processor invokes and performs the operations corresponding to the modules.

In an embodiment, a computer device is provided. The computer device may be a server. An internal structural diagram of the computer device may be shown in FIG. 9. The computer device includes a processor, a memory, and a network interface that are connected by using a system bus. The processor of the computer device is configured to provide computing and control capabilities. The memory of the computer device includes a nonvolatile storage medium and an internal memory. The nonvolatile storage medium stores an operating system, a computer-readable instruction, and a database. The internal memory provides an environment for running of the operating system and the computer-readable instruction in the nonvolatile storage medium. The network interface of the computer device is configured to be communicatively connected to an external terminal via a network. When the computer-readable instruction is executed by the processor, a chatbot configuration method is implemented.

A person skilled in the art may understand that the structure shown in FIG. 9 is merely a block diagram of a part of a structure related to the solutions of this application, and constitutes no limitation on the computer device to which the solutions of this application are applied. Specifically, the computer device may include more or fewer components than those shown in the figure, or combine some components, or have different component arrangements.

A computer device is provided, and includes a memory and one or more processors. The memory stores a computer-readable instruction. When the computer-readable instruction is executed by the one or more processors, the steps of the chatbot configuration method in any one of

the embodiments of this application are implemented.

One or more nonvolatile storage media that store a computer-readable instruction are provided. When the computer-readable instruction is executed by one or more processors, the one or more processors are enabled to implement the steps of the chatbot configuration method in any one of the embodiments of this application.

A person of ordinary skill in the art may understand that all or some of the processes in the methods in the foregoing embodiments may be implemented by a computer-readable instruction instructing related hardware. The computer-readable instruction may be stored in a nonvolatile computer-readable storage medium, and when the computer-readable instruction is executed, the processes in the embodiments of the methods may be implemented. Any reference to the memory, storage, database, or another medium used in the embodiments provided in this application may include a nonvolatile and/or volatile memory. The nonvolatile memory may include a read-only memory (ROM), a programmable ROM (PROM), an electrically programmable ROM (EPROM), an electrically erasable programmable ROM (EEPROM), or a flash memory. The volatile memory may include a random access memory (RAM) or an external cache memory. By way of illustration instead of limitation, the RAM can be in various forms, for example, a static RAM (SRAM), a dynamic RAM (DRAM), a synchronous DRAM (SDRAM), a double data rate SDRAM (DDRSDRAM), an enhanced SDRAM (ESDRAM), a synchlink DRAM (SLDRAM), a Rambus direct RAM (RDRAM), a direct Rambus dynamic RAM (DRDRAM), and a Rambus dynamic RAM (RDRAM).

The technical features of the foregoing embodiments may be randomly combined. For brevity of description, not all possible combinations of the technical features of the foregoing embodiments are described. However, all the combinations of these technical features should be considered to be within the scope of this specification provided that the combinations are not contradictory.

The foregoing embodiments are merely several implementations of this application, and description of the implementations is relatively specific and detailed, but shall not be understood as a limitation on the scope of the present invention. It should be noted that a person of ordinary skill in the art may make several variations and improvements without departing from the concept of this

application, and the variations and improvements shall fall within the protection scope of this application. Therefore, the protection scope of this application patent shall be subject to the appended claims.

CLAIMS

1. A chatbot configuration method, comprising:

obtaining a service form scanning image, extracting form feature information in the service form scanning image, determining a field type of a literal block in the form feature information, identifying, based on the field type, a target literal block whose field needs to be filled in, and establishing an association relationship between literal blocks based on the target literal block whose field needs to be filled in;

determining, by querying a preset data type configuration table based on the target literal block whose field needs to be filled in, a data type of a field that needs to be filled in and that corresponds to each target literal block;

generating a service segment statement of each target literal block based on the data type of the field that needs to be filled in and that corresponds to each target literal block, the association relationship between literal blocks, and a preset statement template; and

configuring a chatbot based on location information of each target literal block in the form feature information and the service segment statement of each target literal block.

2. The method according to claim 1, wherein the obtaining a service form scanning image and extracting form feature information in the service form scanning image comprises:

obtaining the service form scanning image, and preprocessing the service form scanning image;

obtaining location information of each literal block and location information of each literal symbol in the preprocessed service form scanning image based on a trained target detection model, wherein the target detection model is obtained through training by using sample pictures that comprise literal blocks and literal symbols;

dividing the service form scanning image based on the location information of each literal block and the location information of each literal symbol, to obtain multiple literal block images and multiple literal symbol images; and

extracting a literal block and a literal symbol in each literal block image and each literal

symbol image based on a trained picture classification model, to obtain the form feature information in the service form scanning image, wherein the picture classification model is obtained through training by using sample pictures that comprise literal blocks and literal symbols.

3. The method according to claim 2, wherein the obtaining location information of each literal block and location information of each literal symbol in the preprocessed service form scanning image based on a trained target detection model comprises:

inputting the preprocessed service form scanning image into the trained target detection model; extracting features of the preprocessed service form scanning image based on a convolutional neural network in the target detection model; and

obtaining the location information of each literal block and the location information of each literal symbol in the preprocessed service form scanning image based on a fully connected layer in the target detection model and the features of the service form scanning image.

4. The method according to claim 1, wherein the determining a field type of a literal block in the form feature information comprises:

inputting each literal block in the form feature information into a trained classification model, to obtain confidence levels indicating that the literal blocks belong to preset field types;

determining a coordinate distance between each literal block and each literal symbol based on location information of each literal block and location information of each literal symbol in the form feature information;

using each literal symbol whose coordinate distance to each literal block falls within a preset distance threshold range as a literal symbol associated with the literal block;

revising the confidence levels indicating that the literal blocks belong to the preset field types based on an association between each literal block and each literal symbol; and

sorting the confidence levels indicating that the literal blocks belong to the preset field types, and using a field type with the highest confidence level as a field type of each literal block.

5. The method according to claim 1, wherein the field type comprises a field that needs to be filled in, an option field, and a comment field, and the identifying, based on the field type, a

target literal block whose field needs to be filled in and establishing an association relationship between literal blocks based on the target literal block whose field needs to be filled in comprises:

determining, based on a field type of each literal block, that a literal block whose field type is a field that needs to be filled in is the target literal block whose field needs to be filled in;

determining a distance between each target literal block and each option field literal block, and a distance between each target literal block and each comment field literal block based on location information of each literal block in the form feature information;

determining an option field literal block and a comment field literal block corresponding to each target literal block based on the distance between each target literal block and each option field literal block, and the distance between each target literal block and each comment field literal block; and

establishing an association relationship between each target literal block and both the corresponding option field literal block and the corresponding comment field literal block.

6. The method according to claim 1, wherein the configuring a chatbot based on location information of each target literal block in the form feature information and the service segment statement of each target literal block comprises:

determining, based on the location information of each target literal block in the form feature information, a filling sequence of fields that need to be filled in and that correspond to all the target literal blocks;

determining a service process sequence of service segment statements of all the target literal blocks based on the filling sequence of the fields that need to be filled in and that correspond to all the target literal blocks;

generating scenario process information of a service application based on the service process sequence; and

configuring the chatbot by performing model training based on the scenario process information and the service segment statements of all the target literal blocks.

7. The method according to claim 6, wherein the configuring the chatbot by performing model training based on the scenario process information and the service segment statements of all the target literal blocks comprises:

performing model training by inputting the service segment statements of all the target literal blocks into an initial natural language understanding model as a first training set, to obtain a natural language understanding model, wherein the natural language understanding model is used to determine a user intention based on a user statement and capture entity information;

performing model training by inputting the scenario process information into an initial dialogue management model as a second training set, to obtain a dialogue management model, wherein the dialogue management model is used to determine a reply statement based on the user statement and the user intention; and

configuring the chatbot based on the natural language understanding model and the dialogue management model.

8. A chatbot configuration apparatus, comprising:

an obtaining module, configured to obtain a service form scanning image, extract form feature information in the service form scanning image, determine a field type of a literal block in the form feature information, identify, based on the field type, a target literal block whose field needs to be filled in, and establish an association relationship between literal blocks based on the target literal block whose field needs to be filled in;

a first processing module, configured to determine, by querying a preset data type configuration table based on the target literal block whose field needs to be filled in, a data type of a field that needs to be filled in and that corresponds to each target literal block;

a second processing module, configured to generate a service segment statement of each target literal block based on the data type of the field that needs to be filled in and that corresponds to each target literal block, the association relationship between literal blocks, and a preset statement template; and

a configuration module, configured to configure a chatbot based on location information of each target literal block in the form feature information and the service segment statement of each

target literal block.

9. The apparatus according to claim 8, wherein the obtaining module is further configured to obtain the service form scanning image, and preprocess the service form scanning image; obtain location information of each literal block and location information of each literal symbol in the preprocessed service form scanning image based on a trained target detection model, wherein the target detection model is obtained through training by using sample pictures that comprise literal blocks and literal symbols; divide the service form scanning image based on the location information of each literal block and the location information of each literal symbol, to obtain multiple literal block images and multiple literal symbol images; and extract a literal block and a literal symbol in each literal block image and each literal symbol image based on a trained picture classification model, to obtain the form feature information in the service form scanning image, wherein the picture classification model is obtained through training by using sample pictures that comprise literal blocks and literal symbols.

10. A computer device, comprising a memory and one or more processors, wherein the memory stores a computer-readable instruction, and when the computer-readable instruction is executed by the one or more processors, the one or more processors are enabled to perform the following steps:

obtaining a service form scanning image, extracting form feature information in the service form scanning image, determining a field type of a literal block in the form feature information, identifying, based on the field type, a target literal block whose field needs to be filled in, and establishing an association relationship between literal blocks based on the target literal block whose field needs to be filled in;

determining, by querying a preset data type configuration table based on the target literal block whose field needs to be filled in, a data type of a field that needs to be filled in and that corresponds to each target literal block;

generating a service segment statement of each target literal block based on the data type of the field that needs to be filled in and that corresponds to each target literal block, the association relationship between literal blocks, and a preset statement template; and

configuring a chatbot based on location information of each target literal block in the form feature information and the service segment statement of each target literal block.

11. The computer device according to claim 10, wherein when the one or more processors execute the computer-readable instruction, the following steps are further performed:

obtaining the service form scanning image, and preprocessing the service form scanning image;

obtaining location information of each literal block and location information of each literal symbol in the preprocessed service form scanning image based on a trained target detection model, wherein the target detection model is obtained through training by using sample pictures that comprise literal blocks and literal symbols;

dividing the service form scanning image based on the location information of each literal block and the location information of each literal symbol, to obtain multiple literal block images and multiple literal symbol images; and

extracting a literal block and a literal symbol in each literal block image and each literal symbol image based on a trained picture classification model, to obtain the form feature information in the service form scanning image, wherein the picture classification model is obtained through training by using sample pictures that comprise literal blocks and literal symbols.

12. The computer device according to claim 10, wherein when the one or more processors execute the computer-readable instruction, the following steps are further performed:

inputting the preprocessed service form scanning image into the trained target detection model;

extracting features of the preprocessed service form scanning image based on a convolutional neural network in the target detection model; and

obtaining the location information of each literal block and the location information of each literal symbol in the preprocessed service form scanning image based on a fully connected layer in the target detection model and the features of the service form scanning image.

13. The computer device according to claim 10, wherein when the one or more processors execute the computer-readable instruction, the following steps are further performed:

inputting each literal block in the form feature information into a trained classification model,

to obtain confidence levels indicating that the literal blocks belong to preset field types;

determining a coordinate distance between each literal block and each literal symbol based on location information of each literal block and location information of each literal symbol in the form feature information;

using each literal symbol whose coordinate distance to each literal block falls within a preset distance threshold range as a literal symbol associated with the literal block;

revising the confidence levels indicating that the literal blocks belong to the preset field types based on an association between each literal block and each literal symbol; and

sorting the confidence levels indicating that the literal blocks belong to the preset field types, and using a field type with the highest confidence level as a field type of each literal block.

14. The computer device according to claim 10, wherein when the one or more processors execute the computer-readable instruction, the following steps are further performed:

determining, based on a field type of each literal block, that a literal block whose field type is a field that needs to be filled in is the target literal block whose field needs to be filled in;

determining a distance between each target literal block and each option field literal block, and a distance between each target literal block and each comment field literal block based on location information of each literal block in the form feature information;

determining an option field literal block and a comment field literal block corresponding to each target literal block based on the distance between each target literal block and each option field literal block, and the distance between each target literal block and each comment field literal block; and

establishing an association relationship between each target literal block and both the corresponding option field literal block and the corresponding comment field literal block.

15. The computer device according to claim 10, wherein when the one or more processors execute the computer-readable instruction, the following steps are further performed:

determining, based on the location information of each target literal block in the form feature information, a filling sequence of fields that need to be filled in and that correspond to all the target literal blocks;

determining a service process sequence of service segment statements of all the target literal blocks based on the filling sequence of the fields that need to be filled in and that correspond to all the target literal blocks;

generating scenario process information of a service application based on the service process sequence; and

configuring the chatbot by performing model training based on the scenario process information and the service segment statements of all the target literal blocks.

16. One or more nonvolatile computer-readable storage media that store a computer-readable instruction, wherein when the computer-readable instruction is executed by one or more processors, the one or more processors are enabled to perform the following steps:

obtaining a service form scanning image, extracting form feature information in the service form scanning image, determining a field type of a literal block in the form feature information, identifying, based on the field type, a target literal block whose field needs to be filled in, and establishing an association relationship between literal blocks based on the target literal block whose field needs to be filled in;

determining, by querying a preset data type configuration table based on the target literal block whose field needs to be filled in, a data type of a field that needs to be filled in and that corresponds to each target literal block;

generating a service segment statement of each target literal block based on the data type of the field that needs to be filled in and that corresponds to each target literal block, the association relationship between literal blocks, and a preset statement template; and

configuring a chatbot based on location information of each target literal block in the form feature information and the service segment statement of each target literal block.

17. The storage media according to claim 16, wherein when the computer-readable instruction is executed by the one or more processors, the following steps are further performed:

obtaining the service form scanning image, and preprocessing the service form scanning image;

obtaining location information of each literal block and location information of each literal

symbol in the preprocessed service form scanning image based on a trained target detection model, wherein the target detection model is obtained through training by using sample pictures that comprise literal blocks and literal symbols;

dividing the service form scanning image based on the location information of each literal block and the location information of each literal symbol, to obtain multiple literal block images and multiple literal symbol images; and

extracting a literal block and a literal symbol in each literal block image and each literal symbol image based on a trained picture classification model, to obtain the form feature information in the service form scanning image, wherein the picture classification model is obtained through training by using sample pictures that comprise literal blocks and literal symbols.

18. The storage media according to claim 16, wherein when the computer-readable instruction is executed by the one or more processors, the following steps are further performed:

inputting the preprocessed service form scanning image into the trained target detection model;

extracting features of the preprocessed service form scanning image based on a convolutional neural network in the target detection model; and

obtaining the location information of each literal block and the location information of each literal symbol in the preprocessed service form scanning image based on a fully connected layer in the target detection model and the features of the service form scanning image.

19. The storage media according to claim 16, wherein when the computer-readable instruction is executed by the one or more processors, the following steps are further performed:

inputting each literal block in the form feature information into a trained classification model, to obtain confidence levels indicating that the literal blocks belong to preset field types;

determining a coordinate distance between each literal block and each literal symbol based on location information of each literal block and location information of each literal symbol in the form feature information;

using each literal symbol whose coordinate distance to each literal block falls within a preset distance threshold range as a literal symbol associated with the literal block;

revising the confidence levels indicating that the literal blocks belong to the preset field types

based on an association between each literal block and each literal symbol; and

sorting the confidence levels indicating that the literal blocks belong to the preset field types, and using a field type with the highest confidence level as a field type of each literal block.

20. The storage media according to claim 16, wherein when the computer-readable instruction is executed by the one or more processors, the following steps are further performed:

determining, based on a field type of each literal block, that a literal block whose field type is a field that needs to be filled in is the target literal block whose field needs to be filled in;

determining a distance between each target literal block and each option field literal block, and a distance between each target literal block and each comment field literal block based on location information of each literal block in the form feature information;

determining an option field literal block and a comment field literal block corresponding to each target literal block based on the distance between each target literal block and each option field literal block, and the distance between each target literal block and each comment field literal block; and

establishing an association relationship between each target literal block and both the corresponding option field literal block and the corresponding comment field literal block.