Technical Disclosure Commons

Defensive Publications Series

May 2021

AI-ASSISTED GESTURE NAVIGATION FOR COMPUTING DEVICES

Maryam Karimzadehgan

Trystan Upstill

Follow this and additional works at: https://www.tdcommons.org/dpubs_series

Recommended Citation

Karimzadehgan, Maryam and Upstill, Trystan, "AI-ASSISTED GESTURE NAVIGATION FOR COMPUTING DEVICES", Technical Disclosure Commons, (May 03, 2021)
https://www.tdcommons.org/dpubs_series/4262

This work is licensed under a Creative Commons Attribution 4.0 License.
This Article is brought to you for free and open access by Technical Disclosure Commons. It has been accepted for inclusion in Defensive Publications Series by an authorized administrator of Technical Disclosure Commons.

AI-ASSISTED GESTURE NAVIGATION FOR COMPUTING DEVICES

ABSTRACT

A computing device (e.g., a smartphone, a laptop computer, a tablet computer, a smartwatch, etc.) may use a machine learning model to classify user inputs as a back gesture for navigating with respect to graphical user interfaces (GUI) of the computing device. The computing device may apply a machine learning model to input data associated with the user input (e.g., (x,y) coordinates of the user inputs) and a context of the computing device (e.g., an application ("app") that is currently executing on the computing device, the width of the computing device, the orientation of the computing device, etc.) to determine a degree of likelihood of the user input being a back gesture. If the degree of likelihood of the user input being a back gesture satisfies a threshold, the computing device may execute a back action associated with the back gesture. If the degree of likelihood does not satisfy the threshold, the computing device may execute a different action or may discard the user input. The machine learning model may be trained on a computing system (e.g., a remote server) distinct from the computing device while the trained machine learning model may be stored at the computing device.

DESCRIPTION

FIG. 1 below is a conceptual diagram illustrating a system 100 that includes a computing device 102 and a computing system 122. In accordance with various techniques described in this publication, computing device 102 may use a machine learning model 104 to classify user inputs as a back gesture for navigating with respect to graphical user interfaces (GUI) of computing device 100. Machine learning model 104 may be trained on computing system 122 distinct from computing device 102, and computing device 102 may store machine learning model 104.

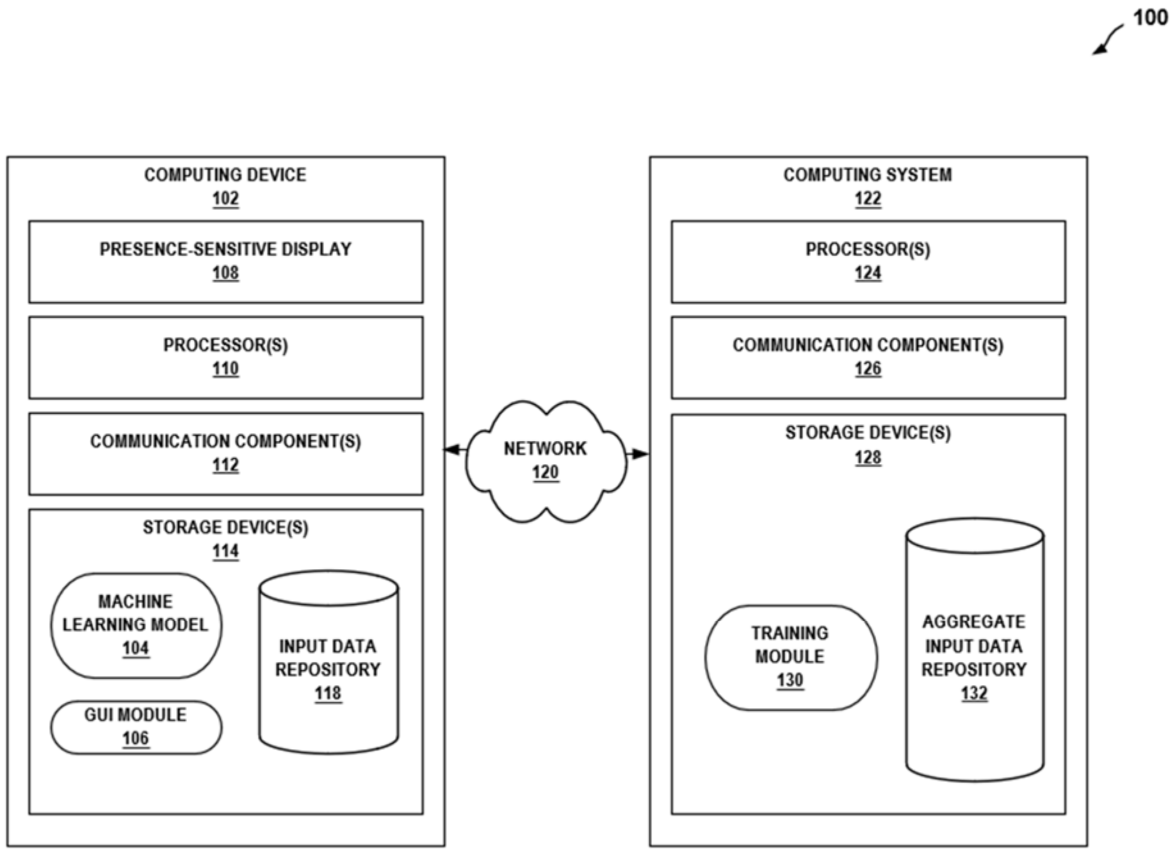

FIG. 1

As shown in FIG. 1, computing device 100 may include a presence-sensitive display 108 (“display 108”), one or more processors 110, one or more communication components 112 (“COMM components 112”), and one or more storage devices 114. Storage devices 114 may include machine learning model 104, a graphical user interface module 106 (“GUI module 106”), and an input data repository 118. Computing device 102 may be any mobile or non-mobile computing device, such as a cellular phone, a smartphone, a desktop computer, a laptop computer, a tablet computer, a portable gaming device, a portable media player, an e-book reader, a watch (including a so-called smartwatch), a gaming controller, and/or the like.

As shown in FIG. 1, computing system 122 may include one or more processors 124, one or more communication components 126, and one or more storage devices 128. Storage devices 128 may include training module 130 and an aggregate input data repository 132. Computing system 122 may be any suitable remote computing system, such as one or more desktop computers, laptop computers, mainframes, servers, cloud computing systems, virtual machines, etc. capable of sending and receiving information via a network 120. In some examples, computing system 122 may represent a cloud computing system that provides one or more services via network 120. For example, computing system 122 may be a distributed computing system.

Display 108 of computing device 102 may be implemented using various display hardware. In some examples, display 108 may function as an input device using a presence-sensitive input component, such as a presence-sensitive screen or touch-sensitive screen, that receives tactile input from a user of computing device 102. The presence-sensitive input component may receive indications of the tactile input by detecting one or more gestures from the user of computing device 102 (e.g., the user touching or pointing to one or more locations of the presence-sensitive input component with a finger or a stylus pen). The presence-sensitive input component may determine a location (e.g., an (x,y) coordinate) of the presence-sensitive input component at which the object was detected. As one example range, the presence-sensitive input component may detect an object, such as a finger or stylus that is within two inches or less of the presence-sensitive input component.

Processors 110 and processors 124 may implement functionality and/or execute instructions associated with computing device 102 and computing system 122, respectively. Examples of processors 110 and processors 124 may include one or more of an application specific integrated circuit (ASIC), a field programmable gate array (FPGA), an application processor, a display controller, an auxiliary processor, a central processing unit (CPU), a graphics processing unit (GPU), one or more sensor hubs, and any other hardware configure to function as a processor, a processing unit, or a processing device. Machine learning model 104, and GUI module 106 may be operable by processors 110 to perform various actions, operations, or functions of computing device 102. Training module 130 may be operable by processors 124 to perform various actions, operations, or functions of computing system 122.

COMM components 112 and COMM components 126 may receive and transmit various types of information, such as input data stored in input data repository 118 and aggregate input data repository 132, over network 120. Network 120 may include a wide-area network such as the Internet, a local-area network (LAN), a personal area network (PAN) (e.g., Bluetooth®), an enterprise network, a wireless network, a cellular network, a telephony network, a Metropolitan area network (e.g., WIFI, WAN, WiMAX, etc.), one or more other types of networks, or a combination of two or more different types of networks (e.g., a combination of a cellular network and the Internet).

COMM components 112 and COMM components 126 may include wireless communication devices capable of transmitting and/or receiving communication signals using network 108, such as a cellular radio, a 3G radio, a 4G radio, a 5G radio, a Bluetooth® radio (or any other PAN radio), an NFC radio, or a WIFI radio (or any other WLAN radio). Additionally or alternatively, COMM components 112 and COMM components 126 may include wired communication devices capable of transmitting and/or receiving communication signals via a direct link over a wired communication medium (e.g., a universal serial bus (“USB”) cable).

Storage devices 114 and storage devices 128 may include one or more computer-readable storage media. For example, storage devices 114 and storage devices 128 may be configured for long-term, as well as short-term storage of information, such as, e.g., instructions, data, or other information used by computing device 102 and computing system 122, respectively. In some examples, storage devices 114 and storage devices 128 may include non-volatile storage elements. Examples of such non-volatile storage elements include magnetic hard discs, optical discs, solid state discs, and/or the like. In other examples, in place of, or in addition to the non-volatile storage elements, storage devices 114 and storage devices 128 may include one or more so-called "temporary" memory devices, meaning that a primary purpose of these devices may not be long-term data storage. For example, the devices may comprise volatile memory devices, meaning that the devices may not maintain stored contents when the devices are not receiving power. Examples of volatile memory devices include random access memories (RAM), dynamic random access memories (DRAM), static random access memories (SRAM), and/or the like.

A user of computing device 102 may provide user input via display 108 to navigate with respect to GUIs of computing device 102. An example user input may include a back gesture (e.g., a swipe from a left edge of display 108 toward a right edge of display 108) associated with a "back" action that causes computing device 102 to navigate from a current GUI to an immediately preceding GUI. However, in general, characteristics of a gesture, including the back gesture, may vary such that computing device 102 may incorrectly process the gesture, potentially resulting in unintended functionality.

For instance, a user, intending to go "back," may perform a back gesture by, for example, swiping from the left edge of display 108 toward the right edge of display 108, but computing device 102 may incorrectly (from the perspective of the user) process the user input as not a back gesture and not go back. In another example, the user, intending to invoke a different action (e.g., a scrolling action, a search action, an edit action, a delete message action, a flag message action, an active application change action, etc.) of computing device 102, such as an application-specific action of swiping to display additional actions of an application, may perform that action's gesture, which may also be a swipe from the left edge of display 108 toward the right edge of display 108. However, computing device 102 may incorrectly process the user input as a back gesture and go back. Such unintended interactions with computing device 102 may be frustrating and inconvenient, potentially harming the user experience of computing device 102.

To increase a probability of computing device 102 correctly processing user inputs for navigating with respect to GUIs of computing device 102, various aspects of the techniques described in this publication enable computing device 102 to use machine learning model 104 to classify user inputs as a back gesture. In some examples, machine learning model 104 may be trained on computing system 122. Training module 130 may train machine learning model 104 on computing system 122 based on training data in aggregate input data repository 132. The training data may include input data that computing system 122 receives from input data repository 118 of computing device 102 (and other computing devices) via network 120. The input data in input data repository 118 may be associated with user inputs provided to computing device 102. For example, user inputs may include the (x,y) coordinates of a gesture from the user of computing device 102, a context of computing device 102 (e.g., the application that is currently executing at computing device 102) when the user performs the gesture, the width of computing device 102, the orientation of computing device 102 when the user performs the gesture, and/or the like.

Training module 130 may train machine learning model 104 by optimizing an objective function. The objective function may represent a loss function that compares (e.g., determines a difference between) output data generated by the model from the training data and labels (e.g., ground-truth labels) associated with the training data. For example, the loss function may evaluate a sum or mean of squared differences between the output data and the labels.

In some examples, training module 130 may train machine learning model 104 using supervised learning techniques. For example, training module 130 may train machine learning model 104 on a training dataset that includes training examples of user inputs labeled as belonging (or not belonging) to the “back gesture” class or “not a back gesture” class. To automate labelling, training module 130 may use heuristics to label the training examples. In some examples, training module 130 may label user input based on subsequent user input indicative of whether the user intended to perform a back gesture or not. In some examples, computing system 122 may update (e.g., periodically) the training dataset used by training module 130 with user input data collected by computing device 102 and other computing devices to further train machine learning model 104.

Responsive to machine learning model 104 being trained, machine learning model 104 may classify a first user input as a back gesture, causing computing device 102 to navigate from a first GUI (e.g., a current GUI) to a second GUI (e.g., an immediately preceding GUI). If a user of computing device 102 provides subsequent user input to return to the first GUI, such subsequent user input may indicate that the first user input was not supposed to be a back gesture, and training module 130 may label the user input accordingly. In another example, machine learning model 104 may classify the first user input as not a back gesture. If the user of computing device 102 provides subsequent user input to cause computing device 102 to navigate from the first GUI to the second GUI, such subsequent user input may indicate that the first user input was supposed to be a back gesture, and training module 130 may label the user input accordingly. In this way, training module 130 may increase the accuracy of machine learning model 104 in classifying user input.

Computing device 102 may download machine learning model 104 (or, if computing device 102 already locally stores a version of machine learning model 104, update machine learning model 104 by downloading a more recent version of machine learning model 104) from computing system 122 and then use machine learning model 104 to classify user input provided to computing device 102. For example, when a user swipes from a left edge of display 108 to a right edge of display (which may, at least, be associated with a back gesture), machine learning model 104 may receive, as input data, the (x,y) coordinates of the user input as well as other input data, such as the context of computing device 102, when the user performs the gesture, the width of computing device 102, the orientation of computing device 102 when the user performs the gesture, and/or the like. Based on these various aspects of the input data, machine learning model 104 may perform binary classification by generating output data that classifies the user input as belonging to either a “back gesture” class or “not a back gesture” class. In some examples, machine learning model 104 may perform classification in which machine learning model 104 provides, for each of the classes, a numerical value descriptive of a degree of likelihood that the input data should be classified into the corresponding class. In some instances, the numerical values provided by machine learning model 104 may be referred to as “confidence scores” that are indicative of a respective confidence associated with classification of the input into the respective class.

In some examples, the confidence scores may be compared to one or more thresholds to render a discrete categorical prediction. For example, machine learning model 104 may generate a confidence score from 0 to 1 for a gesture descriptive of the degree of likelihood of the gesture being a back gesture. In such an example, a confidence score of 0 may be associated with a lowest confidence in the classification and a confidence score of 1 may be associated with a highest confidence in the classification. It should be understood that other ranges of confidence scores (e.g., 0 to 100, 0 to -10, etc.), relationships between confidence scores and amount of confidence (e.g., a confidence score of 0 being associated with a highest confidence and a confidence score of 1 being associated with a lowest confidence), and/or the like are contemplated.

If the confidence score satisfies a threshold, machine learning model 104 may classify the gesture as a back gesture, and computing device 102 may execute the back action. In some examples, the confidence score may satisfy the threshold when the confidence score is equal to or greater than a value of the threshold. For example, if machine learning model 104 generates, for a particular user input, a confidence score of 0.8, and if the value of the threshold for classifying the gesture as a back gesture is 0.7, then machine learning model 104 may classify the particular user input as a back gesture and go back (e.g., to an immediately preceding GUI). On the other hand, if machine learning model 104 generates a confidence score that does not satisfy the threshold, such as a confidence score of 0.6, then machine learning model 104 may classify the particular user input as not a back gesture, and computing device 102 may execute a different action or may discard the user input.

As noted above, computing device 102 may locally execute a version of machine learning model 104 while a version of machine learning model 104 is trained on computing system 122.

In some examples, computing device 102 may update the version of machine learning model 104 stored on computing device 102 by downloading a more recent version of machine learning model 104 (e.g., that has been further trained on computing system 122 based on input data aggregated from various computing devices, including computing device 102). In this way, the accuracy of machine learning model 104 at classifying user input may improve, reducing the frequency of unintended interactions with computing device 102 and potentially improving the user experience of computing device 102.

It is noted that the techniques of this disclosure may be combined with any other suitable technique or combination of techniques. As one example, the techniques of this disclosure may be combined with the techniques described in U.S. Patent Application Publication No. US20180329565A1. In another example, the techniques of this disclosure may be combined with the techniques described in U.S. Patent Application Publication No. US20160018872A1. In yet another example, the techniques of this disclosure may be combined with the techniques described in U.S. Patent Application Publication No. US20130142417A1. In yet another example, the techniques of this disclosure may be combined with the techniques described in U.S. Patent Application Publication No. US20150261298A1.