Assisted digital input for paper forms

Marius Sajgalik

Victor Carbune

Assisted digital input for paper forms

ABSTRACT

Filling out paper forms with various pieces of information is a common task. Filling out the information on paper can be laborious and time-consuming, especially if the form is lengthy. This disclosure describes the use of machine learning and other techniques to assist users in filling out paper forms. The techniques take as input a photo of the paper form and automatically fill out the form to the extent possible with available user information, as permitted by the user. The user can edit the form layout in the digitized version and provide additional input or corrections. A generative model can be used to render the output in printable form with the input data in the style of the form or of the user.

KEYWORDS

- Paper form
- Form layout
- Form structure
- Autofill
- Document editor
- Word processor
- Generative model
- Style transfer

BACKGROUND

Filling out paper forms with various pieces of information is a common task. For instance, paper forms are used for tasks such as opening a bank account, renting an apartment, answering surveys, etc. Filling out the information on paper can be laborious and time-consuming, especially if the form is lengthy.

DESCRIPTION

The described techniques take as input an image of a paper form. For example, the image may be obtained by the user using the camera of a mobile device, a scanner, or an available digital image of a paper form. The obtained image of the form is analyzed to extract the form layout and structure. Fields in the form that are to be filled out are identified. Layout understanding can be performed by a suitable trained machine learning model. The extracted layout is output as bounding boxes that mark form regions along with the probabilities of the corresponding sections requiring input from the user. A user interface is provided to enable the user to refine the system-generated boxes by moving, removing, or resizing. The user can further choose to add boxes in case an input area of the form is not detected automatically.

Optical character recognition (OCR) techniques are utilized to recognize the text present within the form image and map the text requesting user input to corresponding information entity, such as name, date of birth, address, etc. OCR can be performed before or after the form layout analysis, or both tasks can be performed together. If the user permits, user information that is already available from the user's devices or online account is accessed and used to automatically fill in relevant form sections. The user interface enables the user to inspect the pre-filled information and make corrections by editing, deleting, or adding information as needed. If the user denies permission, only such fields that don't require user-specific information are filled in, or an empty form is presented for user input.

For fields in the form that are not filled in automatically, the user can provide the information manually via any suitable and convenient mode such as writing freeform on the device screen, typing characters on a keyboard, speaking the content as voice input, etc. The user input is converted to the appropriate form for the field via appropriate techniques such as handwriting recognition for freeform writing, speech recognition for voice input, etc. The

information is mapped to matching form fields, e.g., by employing a trained machine learning model to rank mappings between the user-entered information and the detected form fields.

Once all required information solicited in the form is completed, a suitable generative model, such as Generative Adversarial Network (GAN), is applied to convert the user input into a format and style suitable for the style of the user and/or the form. For instance, the input can be rendered as printed text that matches the formatting of the form elements. Alternatively, if the user permits, an on-device model can render the input as handwritten text that matches the user's writing characteristics. The digitized form with information filled-out in the appropriate format and style is then rendered in a printable format. Automatically filling out forms in this manner can save users time and effort of filling out the paper form by hand.
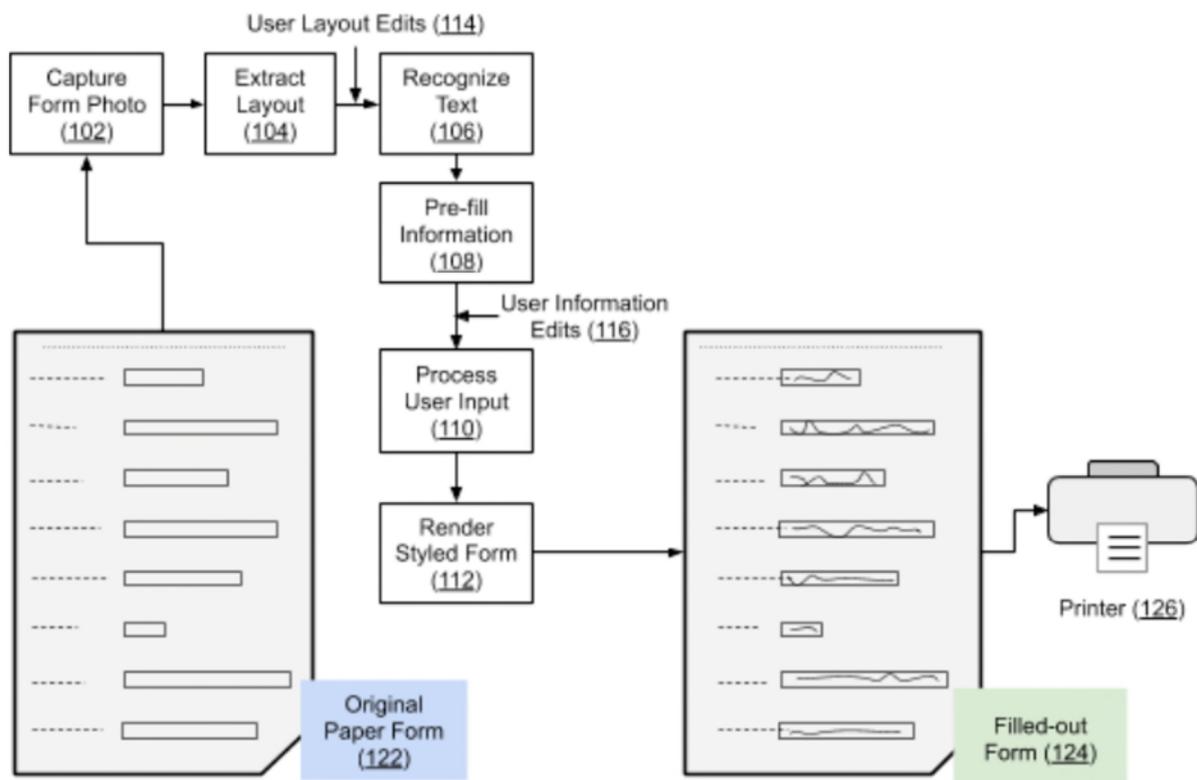

Fig. 1: Digitizing and filling out a paper form

Fig. 1 shows an example process flow of assisted completion of paper forms per the described techniques. A photo (102) of a paper form (122) is captured. The layout and structure of the form is extracted (104) from the photo. After the user performs any needed edits to the extracted layout (114), text within the form is recognized (106), e.g., via OCR. If the user permits, available user information is used to pre-fill corresponding fields in the form (110).

A user interface is provided that enables the user to make edits to the pre-filled information (116) and to provide additional input to complete other portions of the form. As mentioned above, the input can be provided via freeform writing (e.g., on a device touchscreen), keyboard input, voice, etc. The acquired input is converted to text form and mapped to appropriate fields within the form (110). The filled out form is styled to match the desired output format (112) such as handwritten text, printed text, etc. and rendered accordingly as a completed form ready for printing (124). The rendered form can be printed via a printer (126) and submitted in physical form.

The described techniques can be incorporated within document editing applications, image capture/editing applications, etc. Further, additional features such as translating the form elements into another language can also be included.

Further to the descriptions above, a user is provided with controls allowing the user to make an election as to both if and when systems, programs or features described herein may enable collection and/or use of user information. For example, the user can disable automatic form filling or restrict access to certain information. Only such information as permitted by the user is accessed and utilized to automatically complete forms. Thus, the user has control over what information is collected about the user, how that information is used, and what information is provided to the user.

CONCLUSION

This disclosure describes the use of machine learning and other techniques to assist users in filling out paper forms. The techniques take as input a photo of the paper form and automatically fill out the form to the extent possible with available user information, as permitted by the user. The user can edit the form layout in the digitized version and provide additional input or corrections. A generative model can be used to render the output in printable form with the input data in the style of the form or of the user.