

Technical Disclosure Commons

Defensive Publications Series

February 2022

Providing Explicit Shopping Action for Screenshots Containing Purchasable Products

Jing Xiong

Luyao Gong

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

Recommended Citation

Xiong, Jing and Gong, Luyao, "Providing Explicit Shopping Action for Screenshots Containing Purchasable Products", Technical Disclosure Commons, (February 25, 2022)

https://www.tdcommons.org/dpubs_series/4919



This work is licensed under a [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/).

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.

Providing Explicit Shopping Action for Screenshots Containing Purchasable Products

ABSTRACT

Users often take screenshots from shopping websites/apps for later consideration for purchase. However, accessing the product page at a later time requires the user to manually enter the website/product name. This disclosure describes techniques that help users shop for products based on a screenshot of an online page containing information about the product. With user permission, it is determined whether an image on the device is a screenshot. If an image is determined to be a screenshot, a suitably trained machine learning model is employed to determine whether the screenshot contains a purchasable product and OCR is applied to detect whether the screenshot contains text that includes the product price. If a screenshot is identified as containing a purchasable product with a corresponding price, when the user subsequently accesses the screenshot, the shopping action is surfaced using a suitable user interface (UI) mechanism such as a tooltip or button. Interacting with the tooltip or button takes the user directly to the website or app where the product can be purchased.

KEYWORDS

- Screenshot
- Product page
- Product image
- Shopping action
- Shopping reminder
- Shopping intent
- Price information
- Tooltip
- Online shopping
- Action recommendation

BACKGROUND

Users often take screenshots on their device as a reminder to perform an action related to the screenshot. At a later time, users can browse the collection of photos and screenshots on their device to find the screenshot and perform the relevant action. For instance, during the course of browsing a shopping website/app or other activities on a device, users may come across a product they wish to purchase and save a screenshot of the product as a reminder to buy it later. At their convenience, they locate the screenshot of the product in their album. Since the screenshot is not interactive, users must then search separately for the product in the screenshot to find and buy it from an online store.

Many devices and applications have the capability to analyze photos and screenshots using computer vision and/or optical character recognition (OCR). For example, OCR can be applied if a photo or screenshot contains text. The recognized text can then be used in a variety of ways, such as searching the web for the text, translating the text to another language, copy/pasting the text in another application, etc.

However, merely recognizing the text in a screenshot is not sufficient for interactive shopping. Further, if the user's intent is to shop for a product from a screenshot, copying the text from the screenshot is not particularly helpful. Moreover, users are often unaware that text in a screenshot has been recognized via OCR and thus end up searching separately for the product in the screenshot by typing its brand name, description, model number, etc.

DESCRIPTION

This disclosure describes techniques that help users shop for products based on a screenshot of an online page containing information about the product. With user permission, it is determined whether an image on the device is a screenshot. If an image is determined to be a screenshot, a suitably trained machine learning model is employed to determine whether the screenshot contains a purchasable product, such as apparel, furniture, household goods, etc. Next, OCR is applied to detect whether the screenshot contains text that includes the product price.

If a screenshot is identified as containing a purchasable product with a corresponding price, the screenshot is inferred to be connected to an intent to shop for the product. When the user subsequently accesses the screenshot, e.g., while browsing their image library, the shopping action is surfaced using a suitable user interface (UI) mechanism such as a tooltip or button to “Buy the product in this image.” In some examples, the screenshot may be displayed upon capture and the UI mechanism surfaced therewith. Interacting with the tooltip or button takes the user directly to the website or app where the product can be purchased, without requiring the user to perform additional actions such as searching for the product separately, copy-pasting text or URLs, etc.

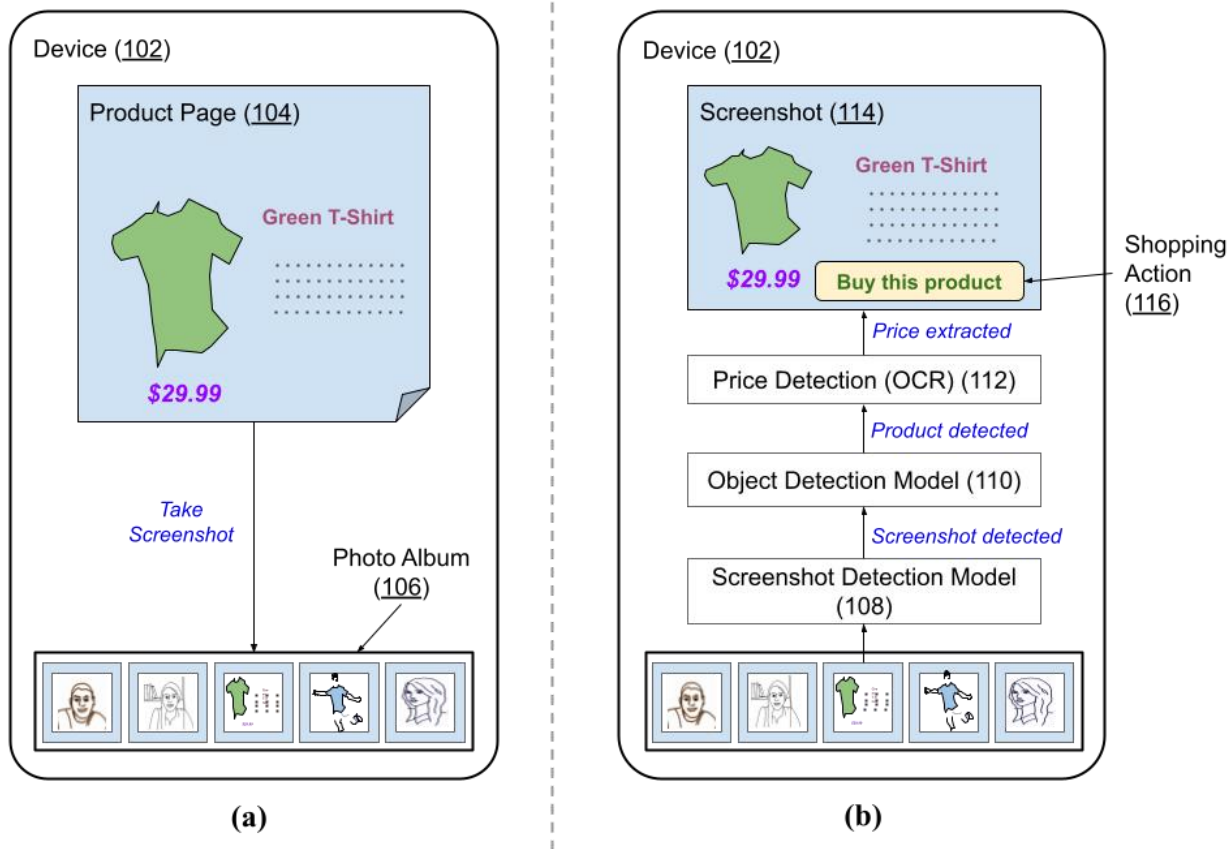


Fig. 1: Detecting a screenshot of a purchasable product and suggesting the shopping action

Fig. 1 shows an example of operational implementation of the techniques described in this disclosure. As shown in Fig. 1(a), a user encounters a product page (104) while using a device (102). The user takes a screenshot of the product page as a reminder to buy it later. The screenshot is saved in the photo album (106) on the device.

As Fig. 1(b) shows, when the user later accesses the screenshot by browsing the photo album, a tooltip/ button with a shopping action (116) is displayed to the user. To generate the shopping action, with user permission, the screenshot is analyzed using a screenshot detection model (108) to determine that it is a screenshot; an object detection model (110) is used to detect that the screenshot contains a purchasable product; and OCR is performed with permission to extract the product price (112). The user can tap on the button with the shopping action, as

displayed overlaid on the screenshot to directly access the product purchase page. Although Fig. 1(b) illustrates a user accessing a screenshot at a later time, in some examples, a screenshot can be analyzed and displayed with a UI element including a shopping action immediately after the screenshot is captured.

Detecting whether an image is a screenshot can be performed by examining the file type of the image. For example, screenshots are typically Portable Network Graphics (PNG) files, unlike photos which are other types of image files, such as Joint Photographic Expert Group (JPEG). Alternatively, or in addition, the image can be input to a machine learning classifier suitably trained to distinguish screenshots from other types of images.

With user permission, the trained machine learning model used to detect whether a screenshot contains a purchasable product can be provided with a labeled list of product categories, such as apparel, furniture, household goods, etc. The presence of a currency symbol along with an amount within the text detected via OCR is an indicator of price information associated with the product detected by the object detector model. Screenshot detection and analysis can be performed either when a screenshot is taken or when the user accesses an image in the image library.

The various trained machine learning models can be run on the user device. Optionally, with user permission, one or more of the models can execute external to the device, such as on a server, within a cloud platform, etc. Any threshold values involved in the model operation can be set by the developers and/or specified by the users and/or determined dynamically at runtime.

With user permission, the techniques described above can be implemented within any device, application, or platform that provides capabilities for capturing, storing, or analyzing screenshots. For instance, the techniques can be integrated within photo library applications, web

browsers, etc. provided on devices such as smartphones, augmented reality (AR) headsets, etc. Provision of a button/tooltip that enables the user to perform a shopping action directly from a screenshot improves the user experience, by eliminating manual steps that are otherwise necessary.

Further to the descriptions above, a user may be 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 of user information (e.g., information about a user's image library, a user's preferences, or a user's current location), and if the user is sent content or communications from a server. In addition, certain data may be treated in one or more ways before it is stored or used, so that personally identifiable information is removed. For example, a user's identity may be treated so that no personally identifiable information can be determined for the user, or a user's geographic location may be generalized where location information is obtained (such as to a city, ZIP code, or state level), so that a particular location of a user cannot be determined. Thus, the user may have 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 techniques that help users shop for products based on a screenshot of an online page containing information about the product. With user permission, it is determined whether an image on the device is a screenshot. If an image is determined to be a screenshot, a suitably trained machine learning model is employed to determine whether the screenshot contains a purchasable product and OCR is applied to detect whether the screenshot contains text that includes the product price. If a screenshot is identified as containing a purchasable product with a corresponding price, when the user subsequently accesses the

screenshot, the shopping action is surfaced using a suitable user interface (UI) mechanism such as a tooltip or button. Interacting with the tooltip or button takes the user directly to the website or app where the product can be purchased.

REFERENCES

1. Goyal, Manoj, Rachit S. Munjal, Sukumar Moharana, Deepak Garg, Debi Prasanna Mohanty, and Siva Prasad Thota. “ScreenSeg: On-Device Screenshot Layout Analysis.” arXiv preprint arXiv:2104.08052 (2021).
2. Cohen, Jason. “How to Take a Screenshot on Any Device” available online at <https://in.pcmag.com/software/43094/how-to-take-a-screenshot-on-any-device> accessed Jan 12, 2022.
3. Harper, Elizabeth. “How to Use Your Smartphone Camera to Search” available online at <https://www.techlicious.com/how-to/how-to-use-your-smartphone-camera-to-search/> accessed Jan 12, 2022.
4. Mrukwa, Grzegorz. “11 Top Image Recognition Apps to Watch in 2021” available online at <https://www.netguru.com/blog/image-recognition-apps> accessed Jan 12, 2022.