October 2022

SHARING THE MAJORITY OF A SCREEN BASED ON OBJECT DETECTION

Yifei Liu

Soya Li

Han Hu

Xueying Xin

Qiaoqian Chen

See next page for additional authors

Inventor(s)
Yifei Liu, Soya Li, Han Hu, Xueying Xin, Qiaoqian Chen, Xiaowei Chen, and Jiaming Ying

SHARING THE MAJORITY OF A SCREEN BASED ON OBJECT DETECTION

AUTHORS:
Yifei Liu
Soya Li
Han Hu
Xueying Xin
Qiaoqian Chen
Xiaowei Chen
Jiaming Ying

ABSTRACT

Users typically do not hesitate to click on a screen sharing button. However, their personal information (such as website bookmarks, notifications, and desktop files) may be revealed unwillingly through such a share. Accordingly, techniques are presented herein that protect a users' private information when they share their entire desktop or separate windows. Aspects of the presented techniques adopt a state of the art (SOTA) object detection model and achieve real-time inference. Further, the incorporated algorithm is lightweight but useful in that it will not impact performance, but it will provide a positive user experience.

DETAILED DESCRIPTION

Users typically do not hesitate to click on a screen sharing button. However, their personal information (such as website bookmarks, notifications, and desktop files) may be revealed unwillingly through such a share.

Accordingly, techniques are presented herein that protect a users' private information when they share their entire desktop or separate windows.

Some existing screen capture solutions can detect the content of a specific layout (such as, for example, a web page within a browser). Such software may read a target window's layout management resource files (such as within, for example, a Qt layout system). However, when it comes to screen sharing such solutions will fail due to a lack of layout management resource files.

Additionally, some existing solutions combine two primary techniques – i.e., screen sharing and majority content detection. Most video conferencing companies develop a portion share mode, which allows users to share only a portion of their screen. However, such a mode must be manually established and typically users begin sharing without a second thought. As a result, such a method is insufficient.

The techniques presented herein protect a user’s private information when the user shares their screens, not only when they share their entire desktop but also when they share individual windows. Users typically do not hesitate to click on a screen sharing button. However, their personal information (such as website bookmarks, notifications, and desktop files) may be revealed unwillingly through such a share. Consequently, aspects of the presented techniques support a sharing mode that smartly proposes the correct screen region that a user may wish to share. Among other things, a user may then tweak the proposed area. For example, if a user is sharing a browser window, then the web page may be displayed on the sharing panel while the bookmarks and the tabs will not be displayed.

Figure 1, below, presents elements of the logic that resides behind aspects of the techniques presented herein.

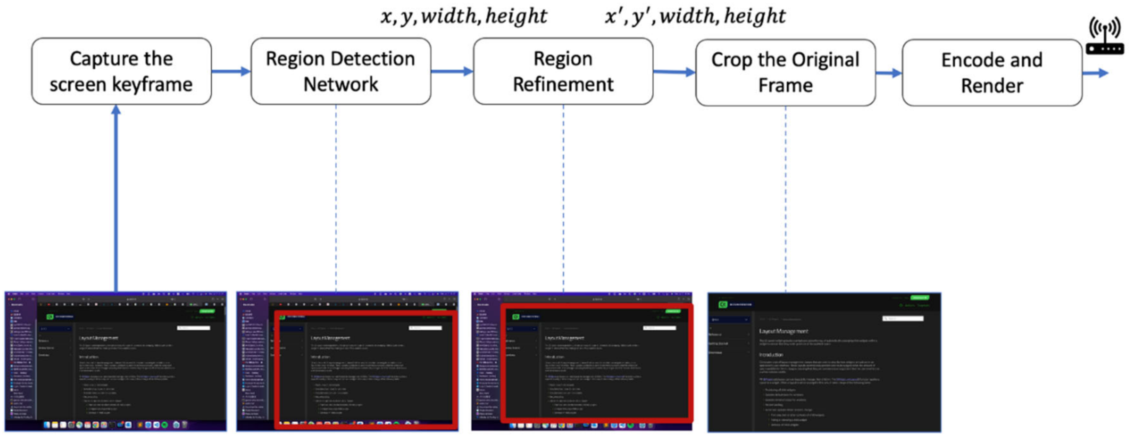


Figure 1: Illustrative Smart Sharing Pipeline

The illustrative pipeline that was presented in Figure 1, above, depicts elements of a smart sharing flow including the actions of capture, region detection, refinement, and encoding.

The techniques presented herein may be further explicated through a description of four key parts of the techniques. A first part of the techniques presented herein encompasses region detection. Under this part, a screenshot keyframe may be fed into a well-trained convolutional neural network (CNN). Such a network may return the top-left coordinates, the width and height of a proposed region, and a confidence score.

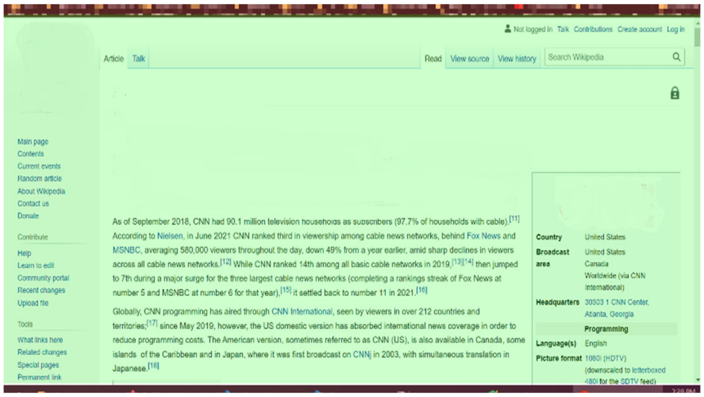
A second part of the techniques presented herein encompasses region rectification involving refining the position of a bounding box. A proposed region might shift slightly from the correct position due to a limitation of the neural network. Thus, the image processing method is designed to refine the box's position (i.e., the coordinates of the top-left corner) by searching the upper corner in a 10 x 10 region around the upper corner of the bounding box and aligning the same with the corner of the highest score.

A third part of the techniques presented herein encompasses cropping the original frame. Under this part, the originally captured keyframe may be cropped based on the box that was identified during the previous steps. Importantly, this region should only contain the main content and the result may be inherited for the following frames until the user changes it.

A fourth part of the techniques presented herein encompasses encoding and rendering. This part incorporates rendering the proposed region and displaying it on the front stage.

Aspects of the techniques presented herein encompass the identification of a proposed region through the application of a CNN. The region proposal model may be developed through a deep learning method. Since this task involves object detection, the state of the art (SOTA) model YOLOv5 may be applied. It is well-known that a good model needs to be driven by well-labeled data. Since there are recordings containing sharing sessions that are available for analysis, it is possible to use those materials to train such a model based on label tools that draw the bounding box of a majority region.

Aspects of the techniques presented herein define two categories of regions. A first category may be referred to as a common area, which consists of the area that is safe for sharing. A second category my be referred to as a sensitive area, which consists of the area containing personal information. Figure 2, below, presents an exemplary screen image that illustrates the above-described region categories.



As of September 2018, CNN had 90.1 million television households as subscribers (97.7% of households with cable).[11] According to Nielsen, in June 2021 CNN ranked third in viewership among cable news networks, behind Fox News and MSNBC, averaging 580,000 viewers throughout the day, down 49% from a year earlier, amid sharp declines in viewers across all cable news networks.[12] While CNN ranked 14th among all basic cable networks in 2019,[13][14] then jumped to 7th during a major surge for the three largest cable news networks (completing a rankings streak of Fox News at number 5 and MSNBC at number 6 for that year),[15] it settled back to number 11 in 2021.[16]

Globally, CNN programming has aired through CNN International, seen by viewers in over 212 countries and territories;[17] since May 2019, however, the US domestic version has absorbed international news coverage in order to reduce programming costs. The American version, sometimes referred to as CNN (US), is also available in Canada, some islands of the Caribbean and in Japan, where it was first broadcast on CNNj in 2003, with simultaneous translation in Japanese.[18]

Country	United States
Broadcast area	United States Canada Worldwide (via CNN International)
Headquarters	30303 1 CNN Center, Atlanta, Georgia
	Programming
Language(s)	English
Picture format	1080i (HDTV) (downscaled to letterboxed 480i for the SDTV feed)

Figure 2: Exemplary Screen Image

In the exemplary screen image that was presented in Figure 2, above, the common area region is depicted in green, and the sensitive area region is depicted in red. Figure 3, below, depicts elements of a corner alignment procedure that is possible according to aspects of the techniques presented herein.

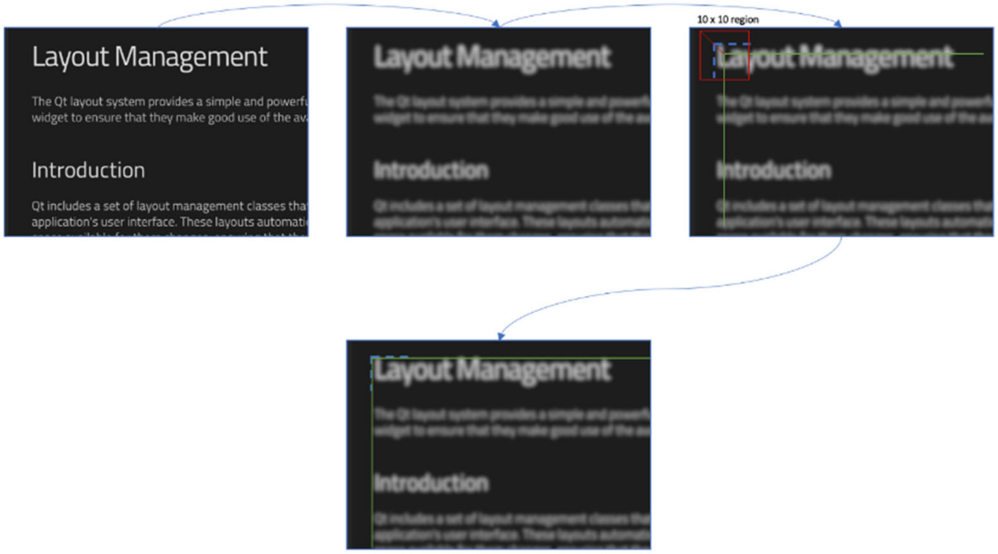

10 x 10 region
Layout Management
The Qt layout system provides a simple and powerf
widget to ensure that they make good use of the av
Introduction
Qt includes a set of layout management classes tha
application's user interface. These layouts automati

Figure 3: Exemplary Corner Alignment Flow

Under the corner alignment procedure that was presented in Figure 3, above, a Gaussian blurred filter may be applied on the y-channel to eliminate noise and subtle sharp edges. Then, the Laplacian operator may be applied around the top-left corner of the proposed region within a 10 x 10 searching window. Finally, a better corner point inside the searching window may be found (as opposed to a CNN-based result) and that point may be selected as the new corner.

Use of the techniques presented herein offers a number of benefits. First, the presented techniques are lightweight. The incorporated algorithm only runs when a user starts a sharing session. Additionally, the algorithm will inherit a proposed region for upcoming frames. As a result, the algorithm will not consume computation resources afterward so it will not impact performance. Second, the presented techniques are secure. For example, a user's personal information (such as email or message preview) will not be revealed during a sharing session. Third, the presented techniques are tolerant of mistakes. For example, if the incorporated algorithm makes a mistake and proposes the wrong region, a user may adjust the proposed region by dragging the indicated rectangle. Fourth, the presented techniques are flexible. In contrast to classic methods, the incorporated algorithm may be adapted to a variety of screen sharing conditions.

In summary, techniques have been presented herein that protect a users' private information when they share their entire desktop or separate windows. Aspects of the presented techniques adopt a SOTA object detection model and achieve real-time inference. Further, the incorporated algorithm is lightweight but useful in that it will not impact performance, but it will provide a positive user experience.