May 06, 2019

Automatic generation of native ad styles using visual attributes of images

Bo Lin

Ravali Pochampally

Yunfan Yang

Automatic generation of native ad styles using visual attributes of images

ABSTRACT

Native advertisements mimic the look and feel of a publisher's content slots and are used to monetize their inventories. Currently, native ad styles are created manually by the publisher based on hand-written rules and heuristics. This can result in ad styles that do not consistently resemble the look-and-feel of the publisher's pages or apps. Also, current techniques generally use the DOM structure or HTML source of the publisher's page or app to generate the native ad. However, the DOM structure or HTML source is not always available, e.g., for apps.

This disclosure describes the use of machine learning techniques to automatically generate native ad styles from key visual attributes of images. The images can be screenshots or design mockups. The techniques can generate native ad styles that match the publisher's look-and-feel closely without recourse to the DOM structure or HTML source for the publisher's page or app.

KEYWORDS

- Native ad
- Content slot
- Visual attribute
- Screenshot
- Design mockup
- Machine learning

BACKGROUND

A native ad is a form of advertisement or other branded content that follows the natural form of the editorial content of a publication. Online publishers, e.g., publishers that target

mobile devices, often, even exclusively, utilize native advertising to monetize their content inventories.

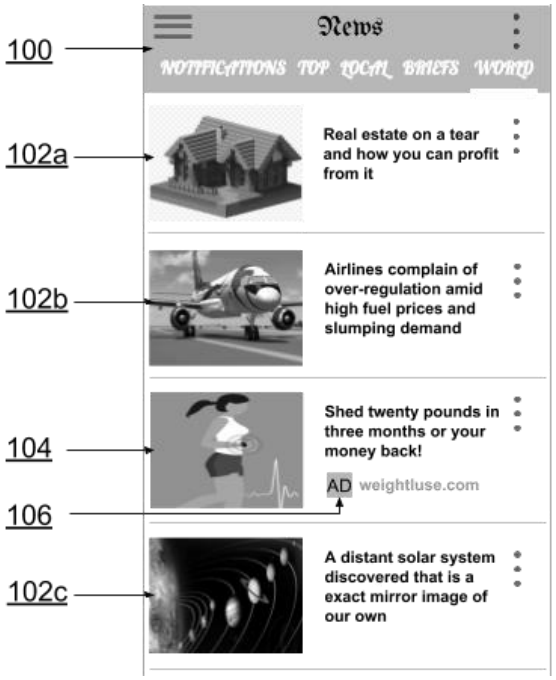

Fig. 1: Example of a native ad

Fig. 1 illustrates an example of a native ad. A publisher provides a news application or website (100) that provides consumers with content (102a-c). Other examples of applications or websites that provide content in the form of a feed are gaming applications, blogging sites, messaging applications, etc. Integrated within the content feed is an advertisement (104). The visual appearance of the advertisement closely mimics editorial content. Such an advertisement is known as a native ad. If the ad is integrated within the feed (as in Fig. 1), it is also known as an in-feed ad. The native ad is marked as such by a marker (106). The marker can also link to the advertiser's website. Native ads appear in many print and electronic publications, and are tailored to the user device, e.g., smartphone, tablet, desktop PC, etc.

For the native ad to perform as designed, e.g., mimic the look-and-feel of editorial content and provide an integrative user experience, publishers provide handwritten rules and

heuristics to ad-networks. Generating native ads using the handwritten rules can result in ad styles that do not consistently resemble the look-and-feel of the publisher's pages or apps. Recent methods of automated generation of native ad styles rely on access to the publisher's document object model (DOM) structure or HTML code, which may not always be available, e.g., for mobile apps.

DESCRIPTION

This disclosure describes the use of machine learning techniques to automatically generate native ad styles from key visual attributes of images. The images can be screenshots or design mockups of the publisher's apps. The techniques can generate native ad styles that match the publisher's look-and-feel closely, without recourse to the DOM structure or HTML source for the publisher's page or app.

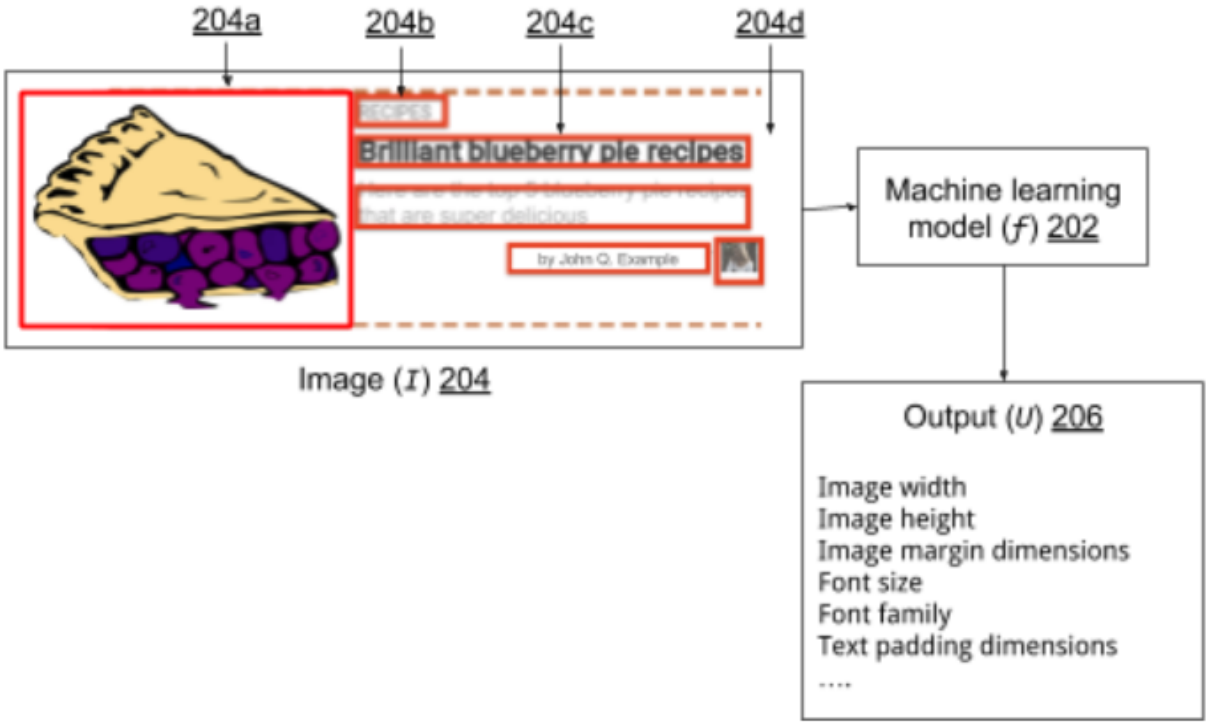

Fig. 2: Automatic generation of native ad styles using visual attributes of images

Fig. 2 illustrates automatic generation of native ad styles using visual attributes of images, per techniques of this disclosure. A machine learning model (202), denoted *f*, accepts as input an image (204), denoted *I*, and generates as output (206) a vector *U* that represents the layout and style of the publisher. The image can include several components, e.g., a sub-image (204a), text (204b-c), margins (204d), etc., each with respective attributes. The machine learning model identifies within its output these components along with their attributes, e.g., image width, image height, image margin dimensions, text font-size, text font-family, text padding dimensions, etc. Mathematically, the map between the output vector *U* and the input image *I* can be represented as

$$f(I) \rightarrow U.$$

A single machine learning model can be used to identify the entire vector *U* of attributes, or separate models can be trained to identify components of the vector. The machine learning model(s) perform one or more of the following actions:

- detect boundaries of the objects in the image, and identify and create masks for the elements of the image;
- map elements of the image, e.g., sub-images, text, etc., into different fields within the native ad specifications;
- extract font information from the text elements;
- extract margin information for all elements;
- combine attributes into the output vector *U*, the set of key visual attributes;
- infer the template, e.g., image left, image right, image top, image bottom, etc. based on the position of the image; etc.

Training the machine learning model

The set *C* of native styles, e.g., native ad elements comprising headline, body, image, etc., with the HTML and the CSS can be generated from an arbitrary visual attribute vector *U*. Mathematically, the map from *U′* to *C* is denoted *g*, such that

$$g(U') \rightarrow C.$$

With the generated native ad style *C*, preview-generation tools, denoted *p*, can be utilized to generate an image *I*:

$$p(C) \rightarrow I.$$

Thus the map between *U′* and *I*, can be expressed mathematically as

$$p(g(U')) \rightarrow I,$$

which is in effect the reverse of the map $f(I) \rightarrow U$ executed by the machine learning model.

To train the machine learning model, the following actions are executed in a loop:

- Generate a random visual attribute vector *U′*.
- Create an image *I* from *U′* using $p(g(U')) \rightarrow I$.
- Use the pair (*I*, *U′*) as training data to train the machine learning model, e.g., train the machine learning model to minimize the loss function

$$\text{loss}(f(I), U').$$

In this manner, a synthetic training data pair (*I*, *U′*) that represents input image *I* and ground truth label *U′* is generated and used to train the machine learning model.

Machine learning techniques can be implemented using deep-learning techniques. The techniques can employ , e.g., neural networks, support vector machines, random forests, boosted decision trees, etc., or other machine learning models that are effective in learning hard-to-describe human intuition in computer vision and understanding problems. One or more different types of neural networks can be used, e.g., LSTM neural networks, recurrent neural networks,

convolutional neural networks, etc. In addition, the machine learning model can be fed with pre-trained image classification feature vectors rather than raw image pixels. Pre-trained image classification feature vectors are high-dimensional transformations of raw image pixels that can speed up image classification. The machine learning model is trained through multiple different placeholder images and texts which can reduce or eliminate bias.

CONCLUSION

This disclosure describes the use of machine learning techniques to automatically generate native ad styles from key visual attributes of images. The images can be screenshots or design mockups. The techniques can generate native ad styles that match the publisher's look-and-feel closely without recourse to the DOM structure or HTML source for the publisher's page or app.