Attribute Adaptation for Personalized Image Search
Adriana Kovashka and Kristen Grauman
(Supplementary Materials)

1 Additional Accuracy Results

In Figure 1, we show more results of the methods' performance as a function of the amount of additional training data beyond D'. This figure is an extended version of Figure 3 in the main paper. We show plots for individual attributes and individual users.

For each dataset, we show 8 cases when our method worked well, and 4 when it failed. In most failure cases, our method was pulled down by inconsistent user responses, as seen by a number of cases where the user-exclusive baseline remains close to chance on binary attributes. Another reason for failure (with respect to the user-exclusive model) were user responses which were the opposite of generic responses (see 'high at the heel' for Shoes Binary in column (f)), where the generic prior can cause negative transfer for our method.

Note that the success of adaptation depends not just on the attribute being learned, but also on individual users. For example, for 'open area' (SUN Binary), our method performs well for one user (column (a)) and poorly for another (column (f)). This motivates one of the future work directions we mention in the main text, namely detecting internal consistency in a user's responses, which should be taken into account when deciding when/how to adapt. Nonetheless, for some attributes like binary 'high at the heel' (column (d)), we see a similar impact for most users.

2 MTurk Qualification Tests

We convey to our Amazon Mechanical Turk workers the basic generic notion of the attributes. We do so to start all users on roughly the same page as far as the meaning intended by the attribute name. While in some cases this could slightly obscure some user perceptual differences (and hence diminish the impact of our method), we determined it would be worthwhile to have sufficient definition of the generic attribute term. Requiring users to complete some very simple qualification tests ensures that the differences in their responses stem from perceptual and not linguistic reasons.

In Figures 3, 4, and 5, we show sample qualification tests. In these tests, the example images or image pairs and the correct answers on the questions asked were determined by hand (for Shoes) or using examples where all labelers agreed (for SUN).

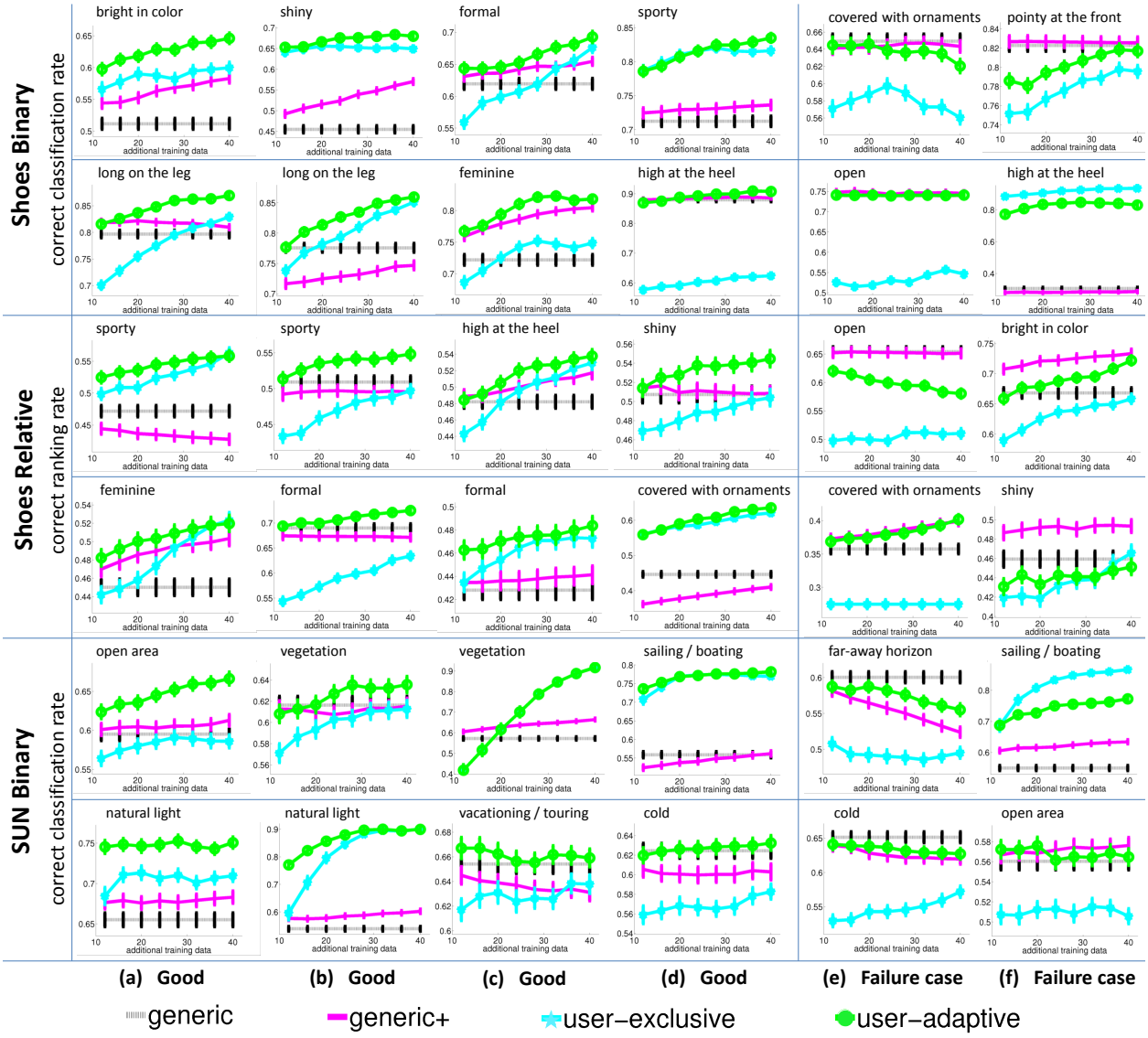


Figure 1: Additional representative results as a function of training data. Note that chance is 50% for binary attributes and 33% for relative attributes. Best viewed in color.

IMPORTANT: Please read the following instructions carefully.
1) Please answer the following questions. You will first need to learn about the meaning of the attributes.
2) When deciding on how to answer these questions, please only base your answers on the information contained in the images.
Thank you!

The shoes in the following images are **pointy at the front**.


The shoes in the following images are **NOT pointy at the front**.


QUESTION:
Is the shoe in the following image
pointy at the front or **NOT pointy at the front**?

ANSWER A:
The shoe is **pointy at the front**.

ANSWER B:
The shoe is **NOT pointy at the front**.

Figure 2: MTurk qualification test for Shoes Binary. Users were required to get 1 of 1 questions right in order to provide user-specific labels.

IMPORTANT: Please read the following instructions carefully.
1) Please answer the following questions. You will first need to learn about the meaning of the attributes.
2) When deciding on how to answer these questions, please only base your answers on the information contained in the images.
3) Please try to give a "more" or "less" response, and reserve the "similarly" response only for cases when it is impossible (or very hard) to make a comparative judgment about the two images because they are too similar in terms of the given attribute.
Thank you!

The shoe in Image 1 is **less pointy at the front** than the shoe in Image 2.

Image 1

Image 2

The shoe in Image 1 is **similarly pointy at the front** as the shoe in Image 2.

Image 1

Image 2

QUESTION:
Is the shoe in Image 1 **more**, **less**, or **similarly pointy at the front** than/as the shoe in Image 2?

Image 1 Image 2

ANSWER A:
The shoe in Image 1 is **more pointy**.

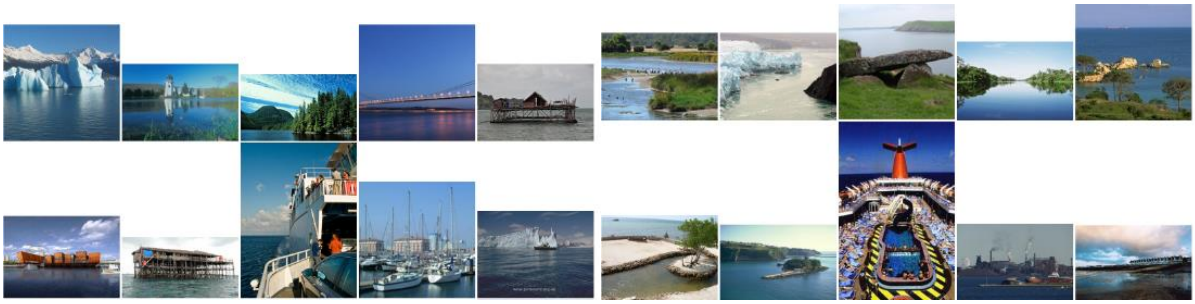
ANSWER B:
The shoe in Image 1 is **less pointy**.

ANSWER C:
The shoe in Image 1 is **similarly pointy**.

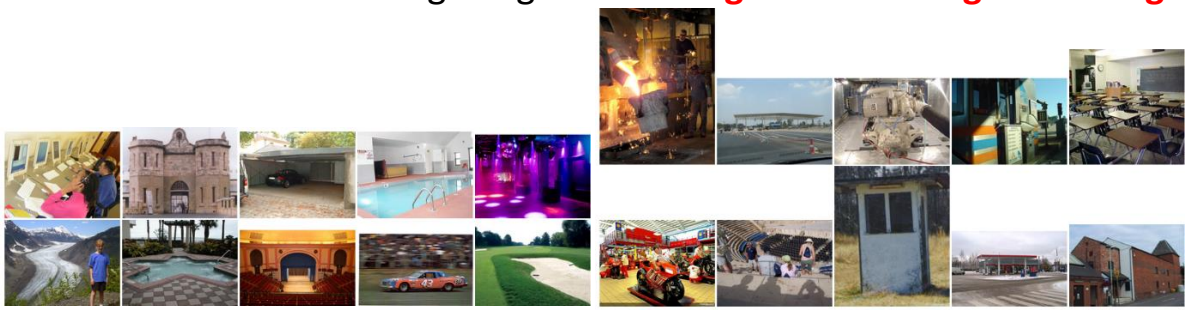
Figure 3: MTurk qualification test for Shoes Relative. Users were required to get 2 of 2 questions right in order to provide user-specific labels.

IMPORTANT: Please read the following instructions carefully.
1) Please answer the following questions. You will first need to learn about the meaning of the attributes.
2) When deciding on how to answer these questions, please only base your answers on the information contained in the images.
Thank you!

The scenes in the following images are **good for sailing or boating**.



The scenes in the following images are **NOT good for sailing or boating**.



QUESTION:
Is the scene in the following image
good for sailing or boating or **NOT good for sailing or boating**?

ANSWER A:
The scene is **good for sailing or boating**.

ANSWER B:
The scene is **NOT good for sailing or boating**.

Figure 4: MTurk qualification test for SUN. Users were required to get 4 of 5 questions right in order to provide user-specific labels.