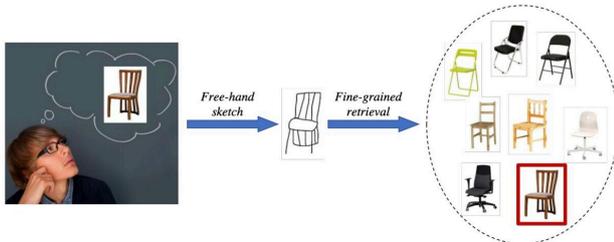# A system to retrieve images using sketches on smart devices

19 July 2022, by Ingrid Fadelli

An illustration of fine-grained sketch-based image retrieval (FG-SBIR), where a free-hand human sketch serves as the query for the instance-level retrieval of images. FG-SBIR is challenging due to 1) the fine-grained and cross-domain nature of the task and 2) free-hand sketches are highly abstract, making fine-grained matching even more difficult. Credit: Bhunia et al.

Researchers at the SketchX, University of Surrey have recently developed a meta learning-based model that allows users to retrieve images of specific items simply by sketching them on a tablet, smartphone, or on other smart devices. This framework was outlined in a paper set to be presented at the European Conference on Computer Vision (ECCV), one of the top three flagship computer vision conferences along with CVPR and ICCV.

"This is the latest along the line of work on 'fine-grained image retrieval,' a problem that my research lab (SketchX, which I direct and founded back in 2012) pioneered back in 2015, with a paper published in CVPR 2015 titled 'Sketch Me That Shoe,'" Yi-Zhe Song, one of the researchers who carried out the study, told TechXplore. "The idea behind our paper is that it is often hard or impossible to conduct image retrieval at a fine-grained level, (e.g., finding a particular type of shoe at Christmas, but not any shoe)."

In the past, some researchers tried to devise

models that can retrieve images based on text or voice descriptions. Text might be easier for users to produce, yet it was found only to work at a coarse level. In other words, it can become ambiguous and ineffective when trying to describe details.

Sketches or doodles, on the other hand, are inherently fine grained and are thus optimal for producing detailed and precise representations of objects. In addition, most modern smart devices have touch screens on which users can produce sketches.

"Key challenges when it comes to sketch-based fine-grained image retrieval are mostly that: (i) people just can't sketch well, (ii) we sketch with different styles and (iii) there are not enough sketches around to train good models," Song explained. "We have published a series of papers on this topic addressing different aspects each time. Our latest paper addresses all three problems at once, and further pushes boundary towards practical deployment of the technology."

The model devised by Song and his colleagues allows even users who are not particularly skilled at sketching to retrieve images of the objects they are seeking, even if it hasn't been trained using images of these objects. This is enabled by its "adaptive" design, which allows the system to adapt to a user's unique drawing style, the quality of his/her drawings and new object categories just using a few example sketches.

Free-hand sketching is ideal for fine-grained instance-level image retrieval. Credit: Bhunia et al.

"Our system learns to work with you (understands your sketches better) very quickly while you are using it for the first few times—typically 2–3 examples are more than enough," the first author, Ayan Bhunia, said. "The best thing is this adaptation happens at testing time only, meaning one does not have to train a new model for a different user/category—this greatly helps practical deployment, just supply the same trained model to every customer and it will learn to work with different style/quality/category once deployed."

In initial evaluations using public datasets, the researchers' model performed remarkably well, as it was able to retrieve images using various sample sketches. In the future, it could be used by online retailers and other companies to allow their customers to find the types of products they are seeking without browsing through their whole catalog.

"Our work is already very mature, the next stage will be to commercialize our system and let ordinary users benefit from this latest development in AI, so that they can find 'that' pair of shoes just by doodling using their fingers on a phone screen," Song added. "In the longer term, we could also extend fine-grained retrieval to the Metaverse. Imagine briefly sketching using your fingers in the 3D world and have the right product/building/object pop up in front of you."

Song and his colleagues are now trying to commercialize their model and promote its introduction in real-world settings. Some world-renowned furniture and clothing retailers have already expressed their interest in using the model to improve their services.

More information: Ayan Kumar Bhunia et al, Adaptive fine-grained sketch-based image retrieval. arXiv:2207.01723v2 [cs.CV], arxiv.org/abs/2207.01723