

Image AI: How brands can benefit from neural networks

DAVID ROSE, MIT MEDIA LAB LECTURER, CEO AND CO-FOUNDER, DITTO LABS 04 APR 2017

A picture says a thousand words, but how can marketers understand what each image means to consumers around the world?

Every image has a meaning. A meaning that extends well beyond a description of what is represented; be it a bunch of flowers, a restaurant, our family or a beach. And these meanings are different, based on our upbringing, culture, and values.

What an 18-year-old woman from Dubai considers ‘luxury’, for example, is radically different to what a middle-aged man from Japan thinks. That’s why the ability to understand and skillfully use images is so vital for brands.

Knowing how people will react to the pictures they use in social media advertising, in their

app. or on the web, will help brands establish more meaningful connections, create more compelling content, and avoid wasting incredible amounts of money by missing their target.

The ability to understand and skillfully use images is vital for brands

Images are the lingua franca

Images are increasingly important in the way consumers talk about themselves, and to each other. As internet analyst Mary Meeker revealed, 3.3 billion photos were shared every day in 2015 on the five biggest western social media platforms — up from just under two billion in 2014. Include the Chinese internet players and the number would be even higher.

Meeker also notes that Gen Z — those born up to 20 years ago — communicate primarily with images, compared to millennials who communicate with text. Her 2016 presentation predicts that in five years at least 50 % of all searches are going to be images or speech.

Already, Google Photos, and Apple iOS 10 lets us search for objects in photos. How we capture images is also changing dramatically. Cameras are appearing everywhere; inside fridges to reorder food, inside jewelry and watches to help us take selfies, and inside sunglasses and tie-clips to record snippets of video.

The photo remix society

Consumers are also becoming better at manipulating pixels. While this may be the age of the selfie — 30 % of photos contain faces — it's also rapidly becoming the age of the photo editor.

Many have become skilled at editing on the move by cropping, colourising and adding

filters. Over 5 % of photos now include superimposed text as people create more memes, overlays, grids and other remixes. Pristine original photos are a thing of the past.

While this may be the age of the selfie — 30 % of photos contain faces — it's also rapidly becoming the age of the photo editor

As cameras become ubiquitous in malls and cities, it raises the question about how the visual web might turn into the big brother web. Belgium digital artist Dries Depoorter, who specialises in challenging surveillance and privacy concerns, has an artistic response to this consideration.

Depoorter's 2016 installation, 'Jaywalking', allows onlookers to view traffic webcams and decide on the fate of pedestrians recklessly crossing the road in a town in Canada. Viewers can press a special button to email a screen shot of the violation to the nearest police station.

The rise of machine learning

It's impossible for marketers to sift through the billions of social media images shared every day to understand the meaning and context of brand images. Thankfully, machine learning is making this easier by helping us identify cultural insights coded into images.

Artificial Intelligence (AI) is experiencing a renaissance where powerful cloudbased servers are able to automatically 'understand' the content of a photo or video. The industry calls this 'image understanding'. It includes the ability to locate brands, objects, context, and recently more subjective qualities of the photo.

Machine learning is making it easier to identify cultural insights coded into images

That's what we do at Ditto Labs. We read the torrent of photos shared on social media to understand how people use brands 'in the wild' and track these trends. Our Visual Brand Power score allows agencies and advertisers to see what is trending in social media using image recognition.

Recently, we have built a set of deep neural networks to discover qualities of photos that would previously have been the exclusive providence of a human art director. Our neural networks can score photos on anything, from identifiers like 'romantic' and 'modern', to 'alluring' and 'luxurious'.



Above: What 'luxury' looks like in the UAE

How a neural network works

Mimicking the neurophysiology of the human brain, our computers are able to learn the

signals of a given concept by ‘seeing’ many examples. These examples are called training data. Given enough training data, neural networks can learn to recognise almost anything that humans can.

The advantages that computers provide, of course, are speed and cost. A neural network can be spread across hundreds of computers to ‘read’ hundreds of millions of photos per day for about the cost of a single (well-paid) computer programmer in Boston.

As an example, to find training data that represents the concept of luxury, our team at Ditto used a two-step process. First, we automatically searched Twitter data to find photos that were labelled explicitly with #luxury. We found 18,000 photos with this hashtag over a few months.

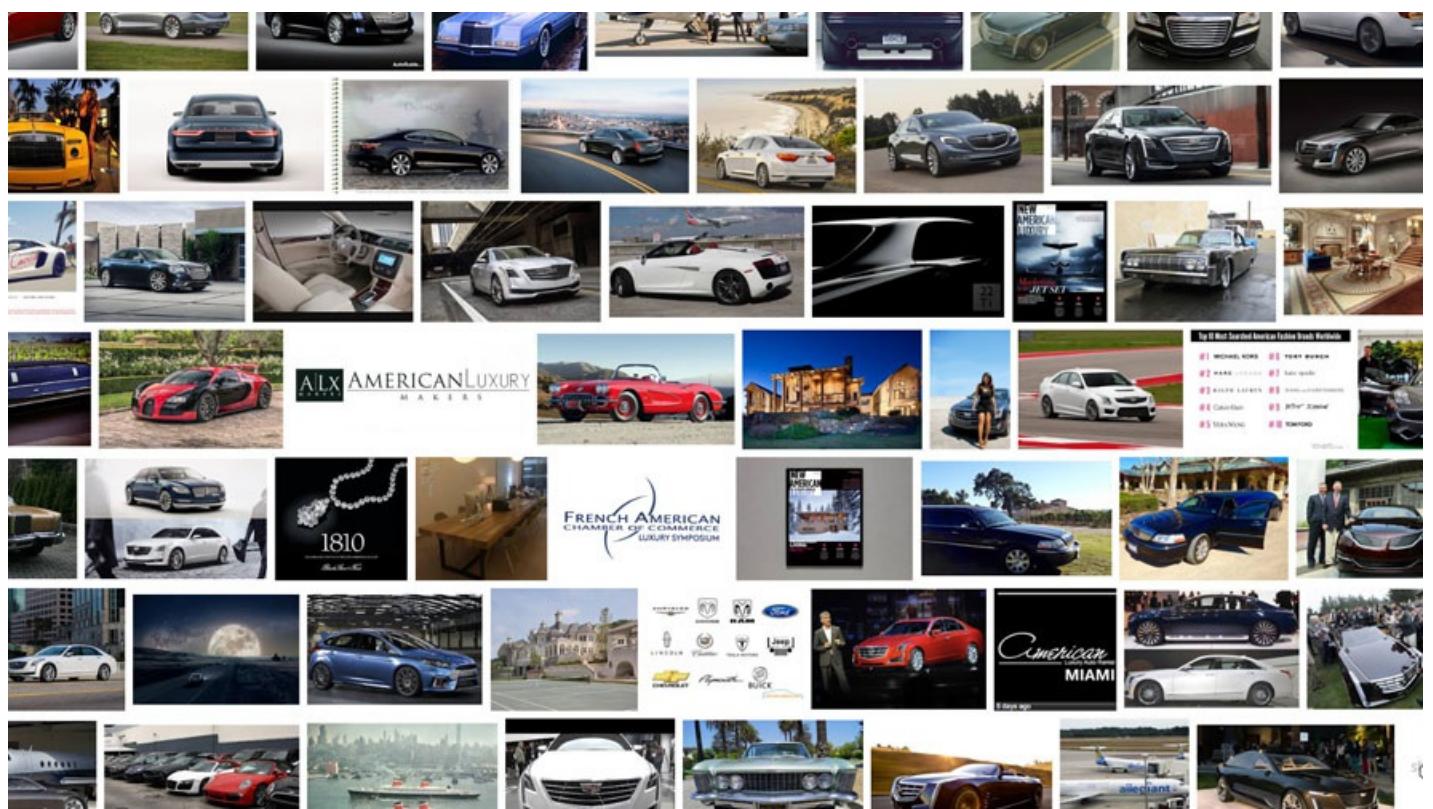
Marketers can use neural networks to create mood boards of images around certain keywords

Second, we asked hundreds of real people from different cultures to manually select the photos from this set that represented their own personal concept of luxury. As we found, someone from New York, for instance, might have a very different concept of ‘luxury’ to someone in Dubai.

This helped us compile 1,471 examples of positive matches (images that the majority of people agreed represented ‘luxury’) and 1,485 negative matches (images where opinion was split). We fed this data into an internal software tool to build a neural network capable of ‘classifying’ or identifying other images people from specific countries would consider ‘luxury’.

We kept 20 % of our training set in reserve to test the suggestions from our neural network.

When the network was asked to rate a photo previously labelled ‘luxury’ by a human, it agreed 98.4 % of the time.



Above: What ‘luxury’ looks like in the USA

Automatically personalising images for different audiences

This level of accuracy creates real opportunities for brands looking to share culturally relevant images with target audiences all around the world.

Firstly, they can use neural networks to create mood boards of images around certain keywords or topics like ‘family friendly’ or ‘luxury’. Brands can use this information to tweak their campaign visuals to cater to local tastes.

Secondly, neural networks can help brands deliver more personalised messages, and help them reveal the right images to the right consumers. At Ditto, we helped one of our clients improve engagement by 20 % by sharing photos matched against their audiences’ cultural aspirations.

Finally, image awareness gives brands a new source of creative content, allowing them to

populate their social media activity with more relevant and appealing images. Get the images right, and they'll not only inspire their followers but encourage their followers' friends to learn more or shop for deals too.

Since the advent of photography, consumers have learned how to instinctively interpret images every day. This skill can now be mimicked by a computer and be conducted at scale. Machine learning and neural networks make the global photo album accessible and understandable and enables marketers to ensure their messages remain relevant to their consumers' needs.

Author's Note: Ditto Labs is a pioneer of 'visual listening'. Its proprietary image recognition software discovers and tags brands, scenes, and objects in online photos.

Get insights in your inbox

Interested in the future of media and marketing?

Get actionable advice, insights and opinion from MediaCom experts every month

[Sign me up](#)
