New Methods of Image Search

Kai Uwe Barthel

HTW Berlin, University of Applied Sciences
visual-computing.com

pixolution.org

pixolution

htw
Hochschule für Technik
und Wirtschaft Berlin

University of Applied Sciences

Summary

- History and status quo of image search
- Machine learning, Deep learning
  AI, Convolutional networks, ...
  started a new era of image search:
  - Automatic tagging
  - New image descriptors
  - New image exploration tools
  - New image manipulation tools &
    image synthesis

Image Search:
Status Quo
&
Previous Approaches

There are too many images

- Constantly growing number of images
- Image search is very time consuming
- Only few images can be looked at simultaneously
- No visual browsing schemes are available

Only few images can be looked at at once

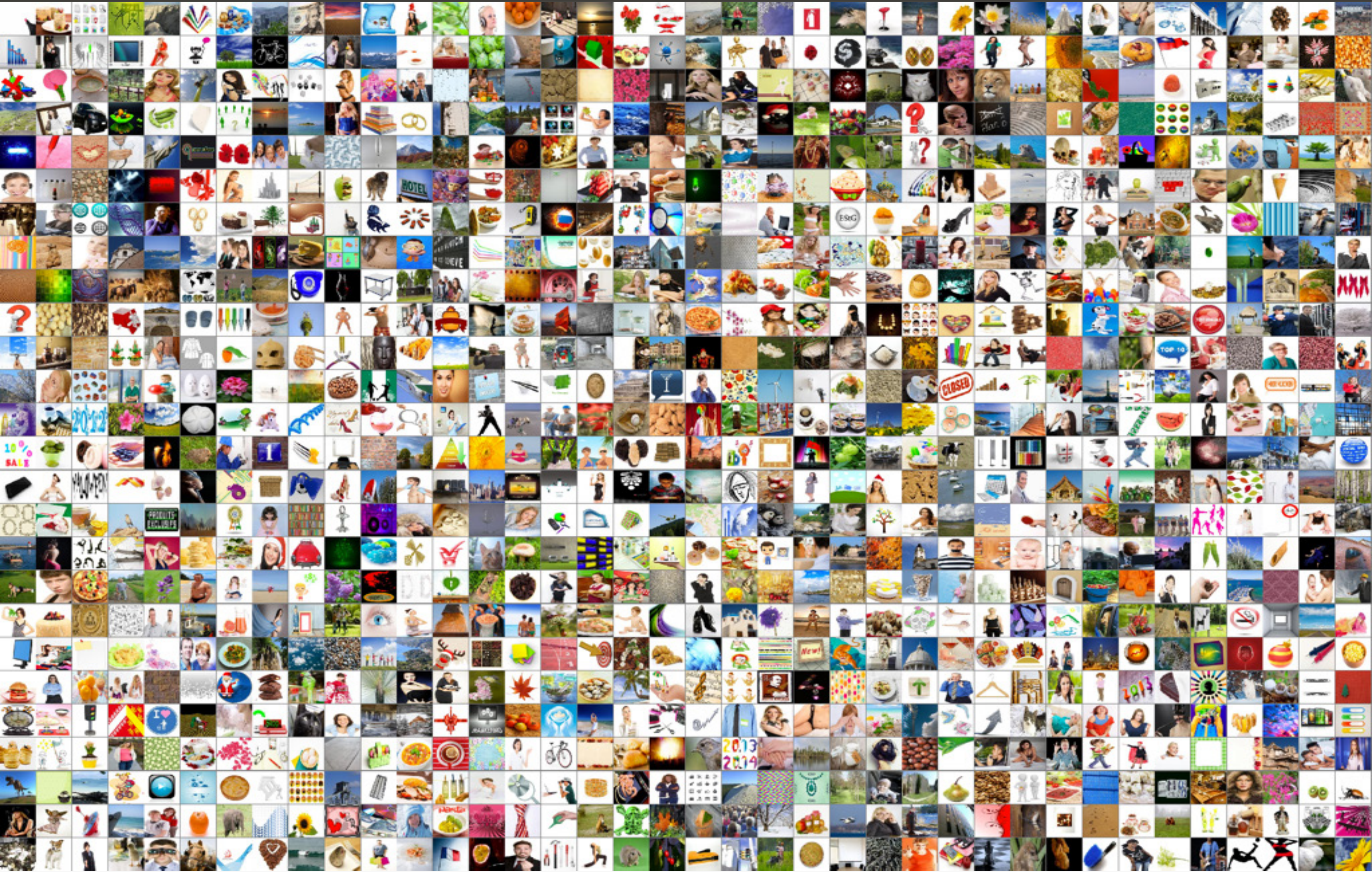
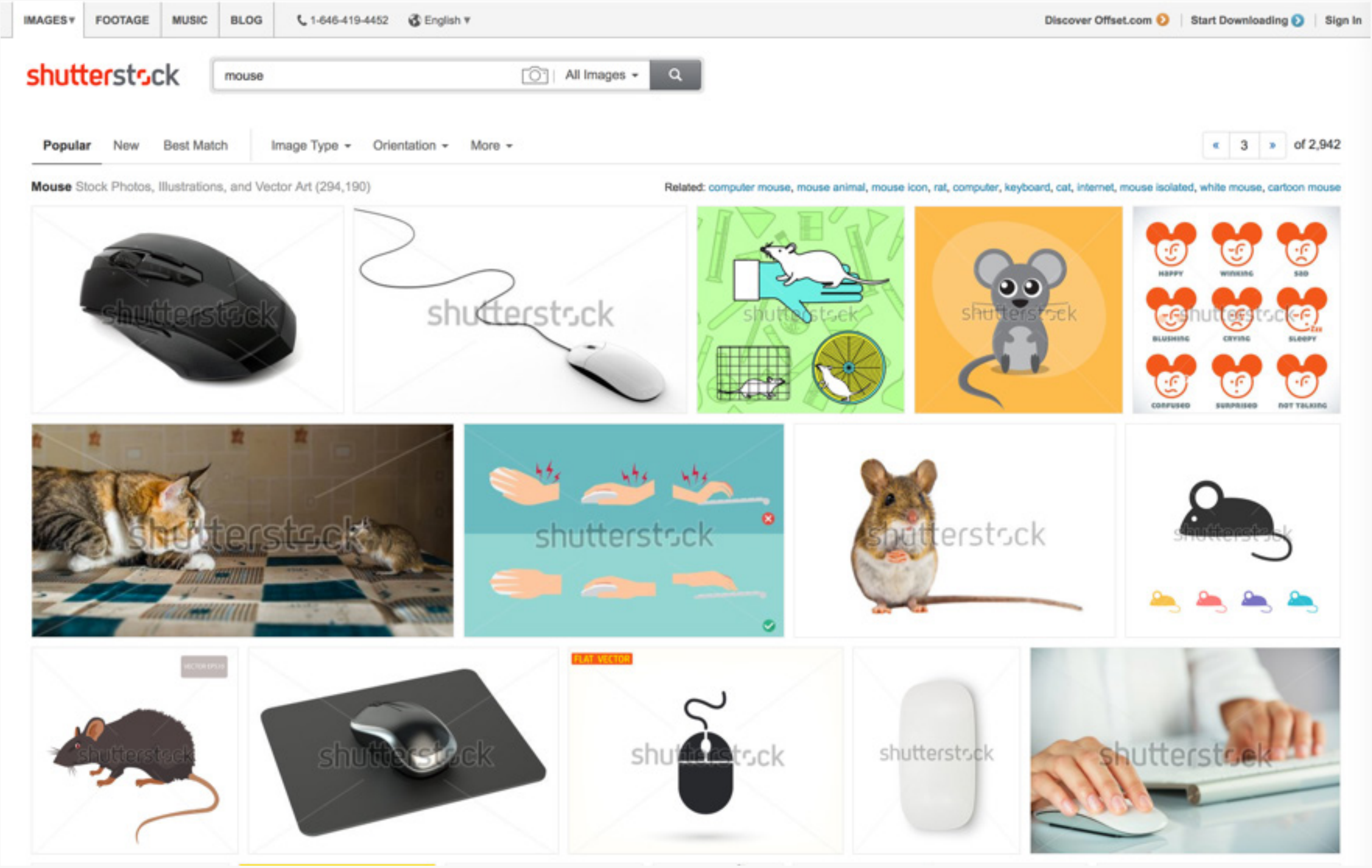


Image search as we know it …

Keyword search



Keyword search

- Good keywording is expensive, keywords often are incomplete, overloaded or wrongly translated
- Exact keyword matching leads to finding everything or nothing

meadow

826,951 results:

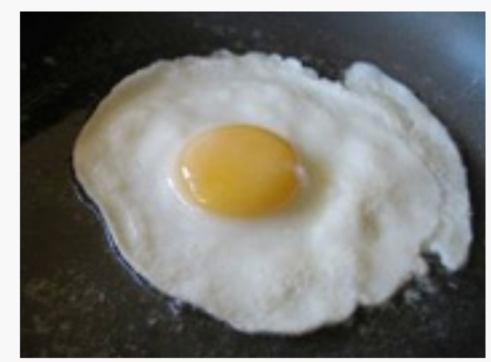
meadow blue sky apple tree clouds flowers family

3 results:

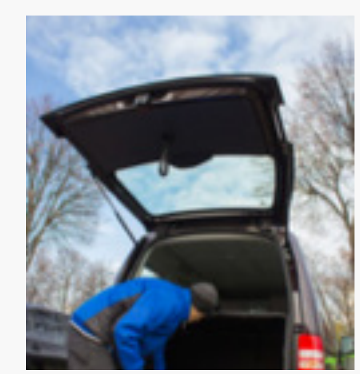
Low-Level Content based Image Search

Search for images with similar colors, textures or shapes

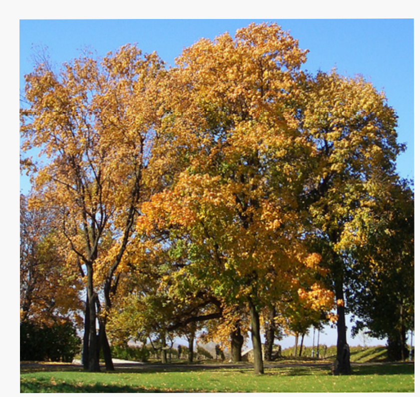

- Finds similar looking images.




- Does not understand the meaning of the images.

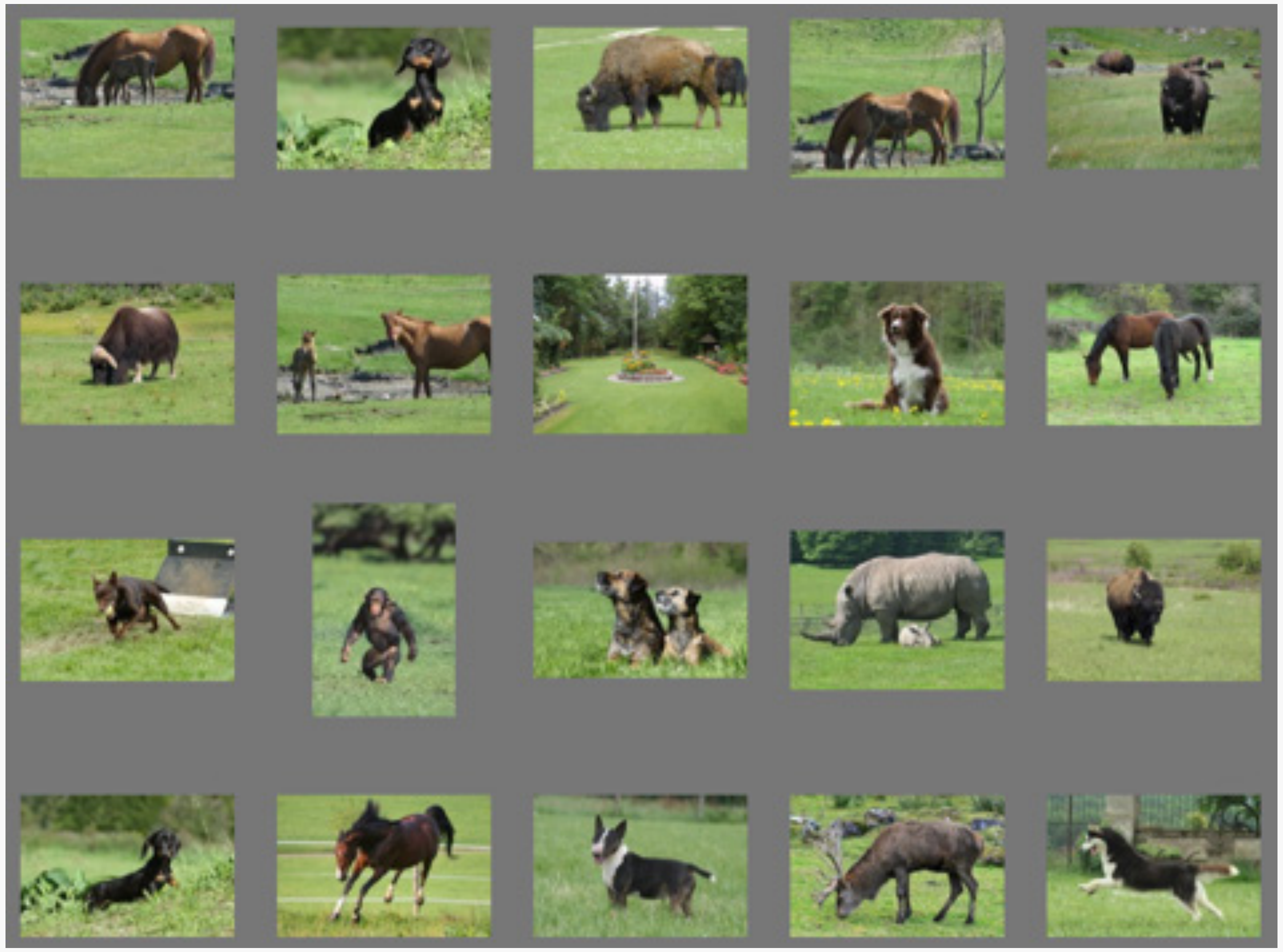



- Cannot find similar images that look different




Problem of low-level visual search

Query: 

Result: 

Visually similar images,

content may be different

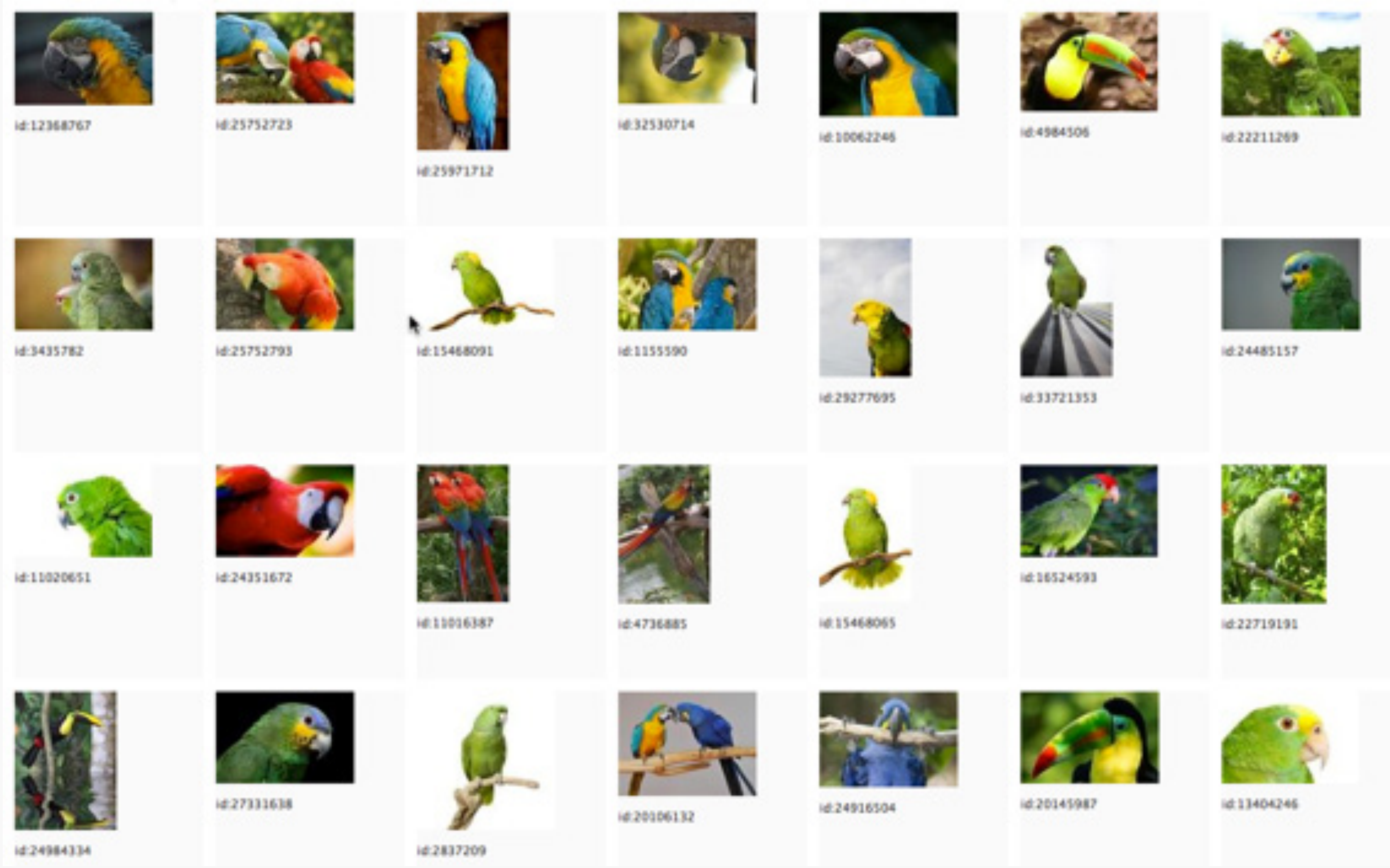
"More like this" keyword search

Query:

america amazon animal beak bird brazil era eye feather fuss parrot portrait yellow ...

(all keywords of the query image)

Result:



Similar image content,

large appearance variations

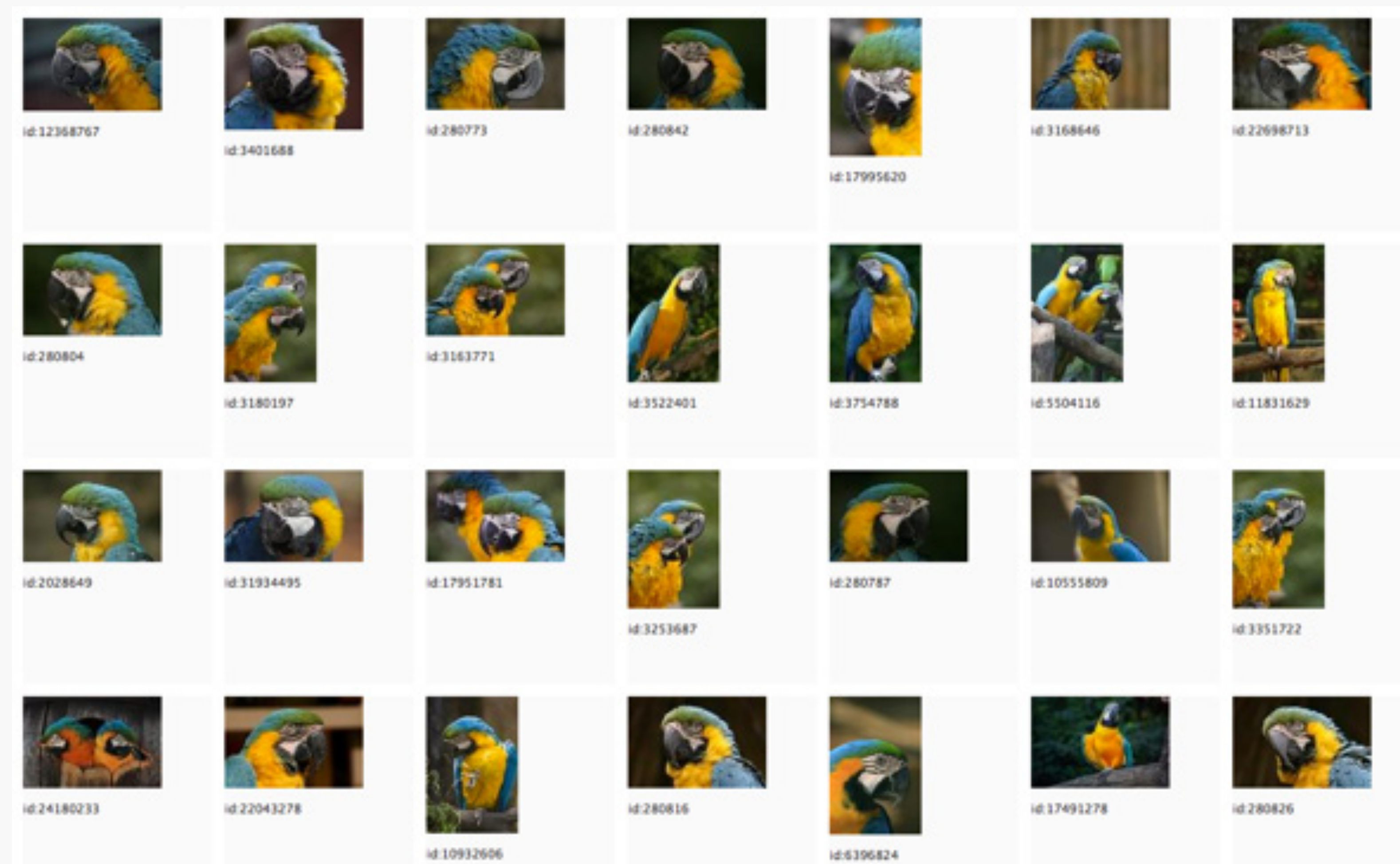
Fusing visual and keyword search

Query:

+

america
amazon animal
beak bird
brazil era eye
feather fuss
parrot portrait
yellow ...

Result:



- Visually similar images and
- Similar content

But …

Needs for finding images more easily

1. Most images are untagged

→ need for
automatic image understanding

2. There are far too many images

→ need for
visual image browsing schemes

Automatic image understanding using AI, ML, Deep Learning

XKCD September 2014

HTTP://XKCD.COM/1425/

In the 60s, Marvin Minsky assigned a couple of undergrads to spend the summer programming a computer to use a camera to identify objects in a scene.
He figured they'd have the problem solved by the end of the summer.
Half a century later, we're still working on it.

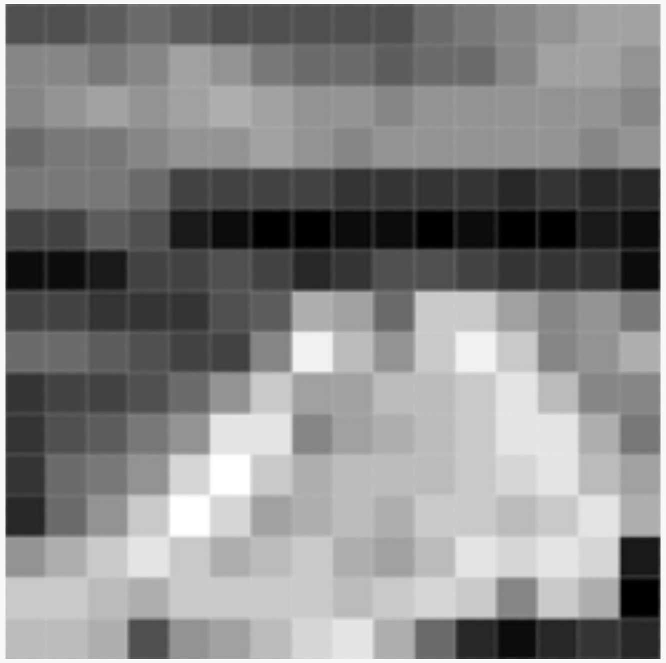
Today only 2.5 years later
the problem is solved

Automatic image understanding

Automatic image understanding

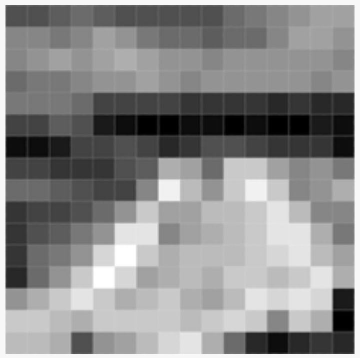
3264x2448 pixels
16,7 millions
of colors

16x16 pixels



16x16 pixels
16 brightness
levels

Automatic image understanding



16x16 pixels, 16 brightness levels

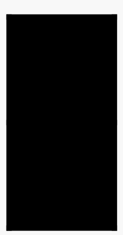
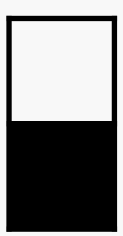
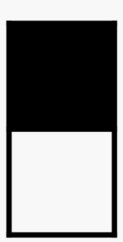
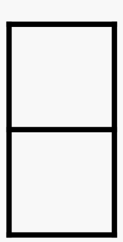
The number of possible images:

16^{16x16} =

179 769 313 486 231 590 772 930 519 078 902 473 361 797 697 894 230 657 273 430 ⋱
081 157 732 675 805 500 963 132 708 477 322 407 536 021 120 113 879 871 393 357 ⋱
658 789 768 814 416 622 492 847 430 639 474 124 377 767 893 424 865 485 276 302 ⋱
219 601 246 094 119 453 082 952 085 005 768 838 150 682 342 462 881 473 913 110 ⋱
540 827 237 163 350 510 684 586 298 239 947 245 938 479 716 304 835 356 329 624 ⋱
224 137 216

Automatic image understanding

- Easy case: Image with 2 pixels, black&white
- Possible questions: Are both pixels ...

	black?	white?	different?
	yes	no	no
	no	no	yes
	no	no	yes
	no	yes	no

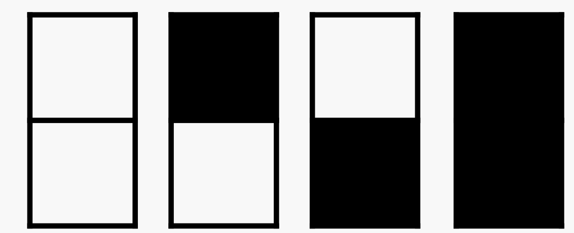
Neuronal Networks (NN)

- Origin in the 60s (Perceptron)
  Impossible to solve the XOR problem

- Restart in the 80s:
  Hidden layers & Backpropagation
  Not suited for image recognition

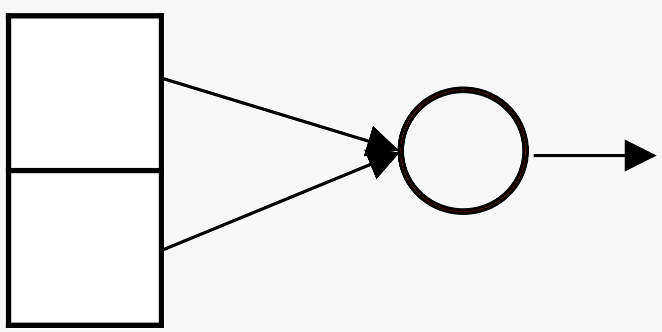
- Restart a few years ago:
  Today the best technology for
  solving all kind of problems for which
  humans cannot describe an algorithm

Neural Networks (NN)

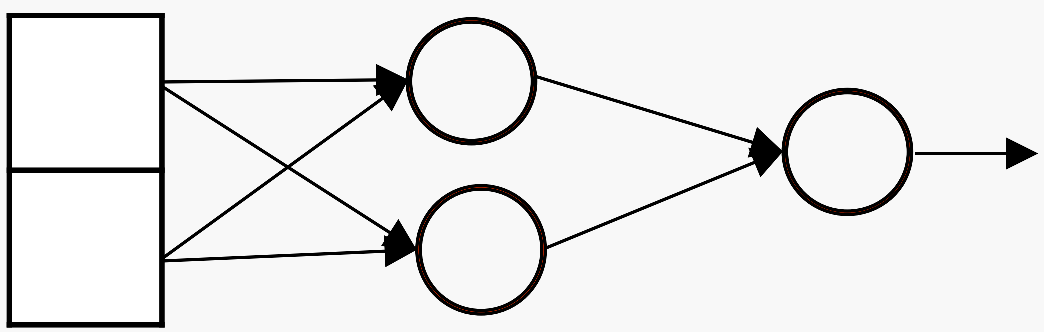


…can be trained to answer these questions:

Are both pixels white?

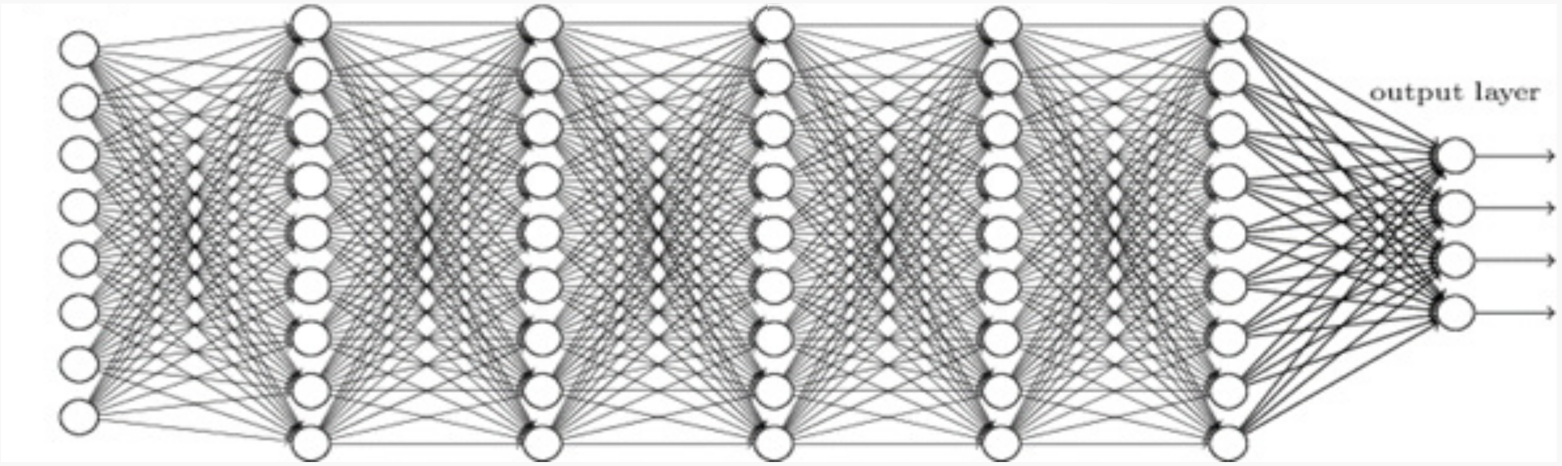


Are both pixels different?



Idea: Use huge networks to make image recognition possible …

Image recognition with NN

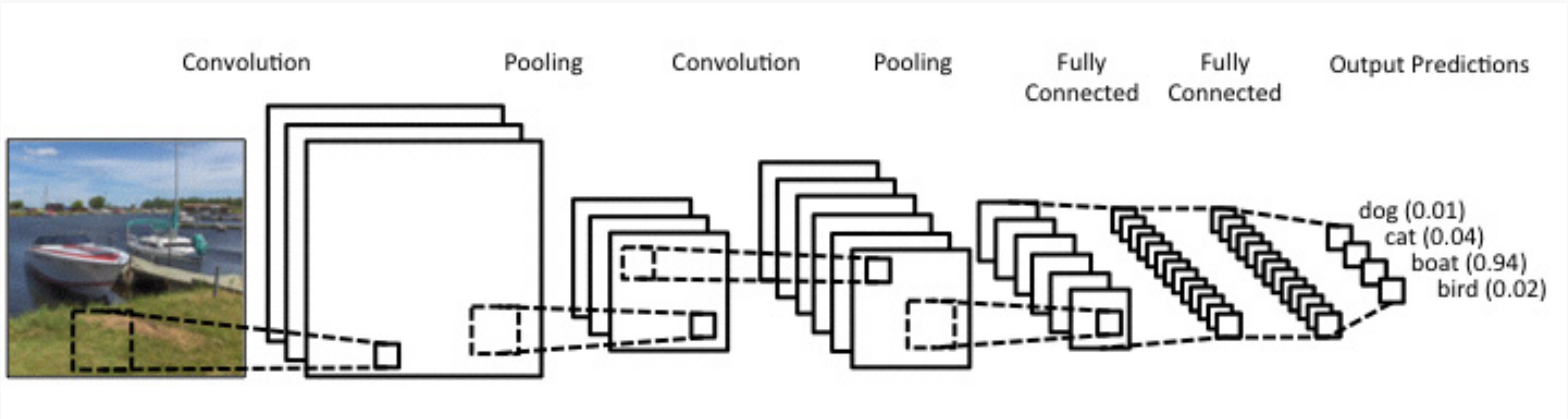


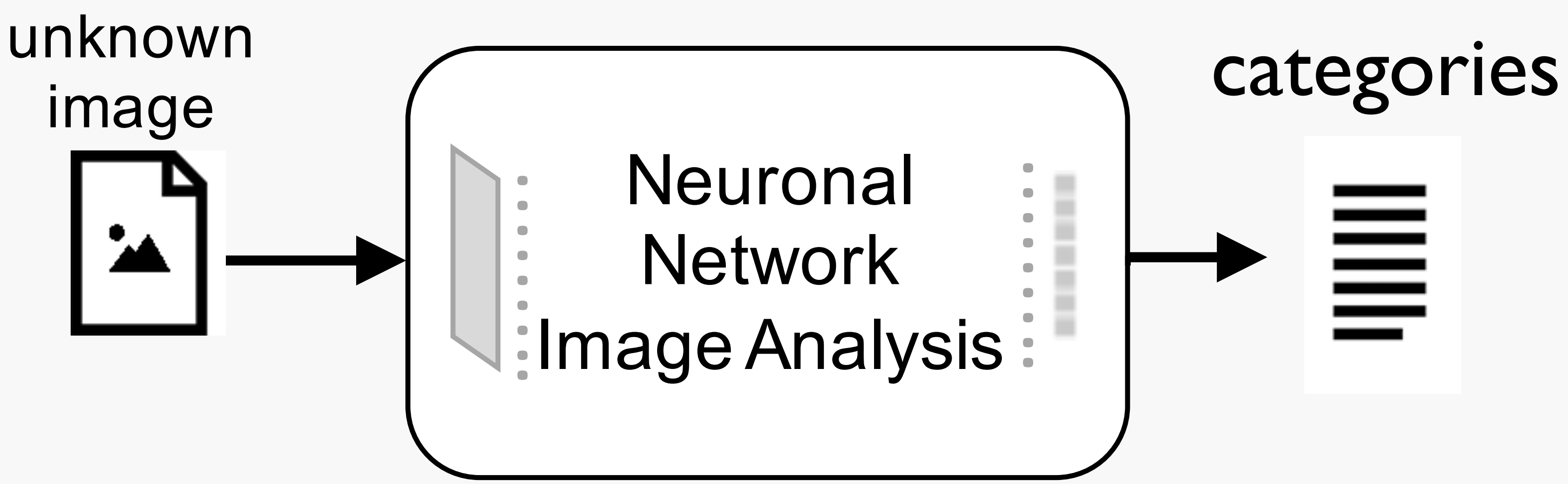
dog
cat
mouse
car
...

Did not work for a long time! :(

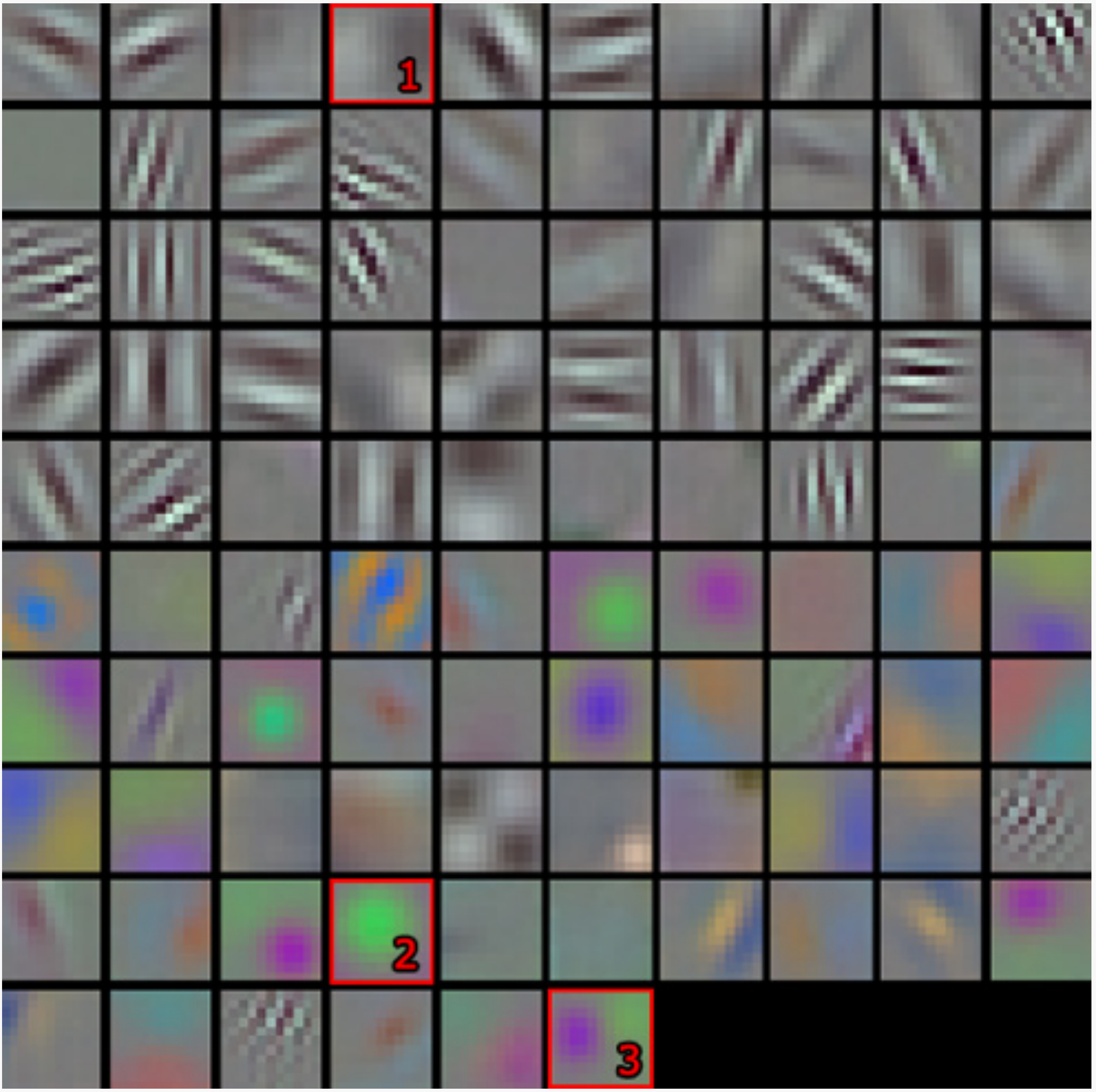
- Not enough tagged example images
- Computers were too slow
- Fully connected neural networks were hard to train

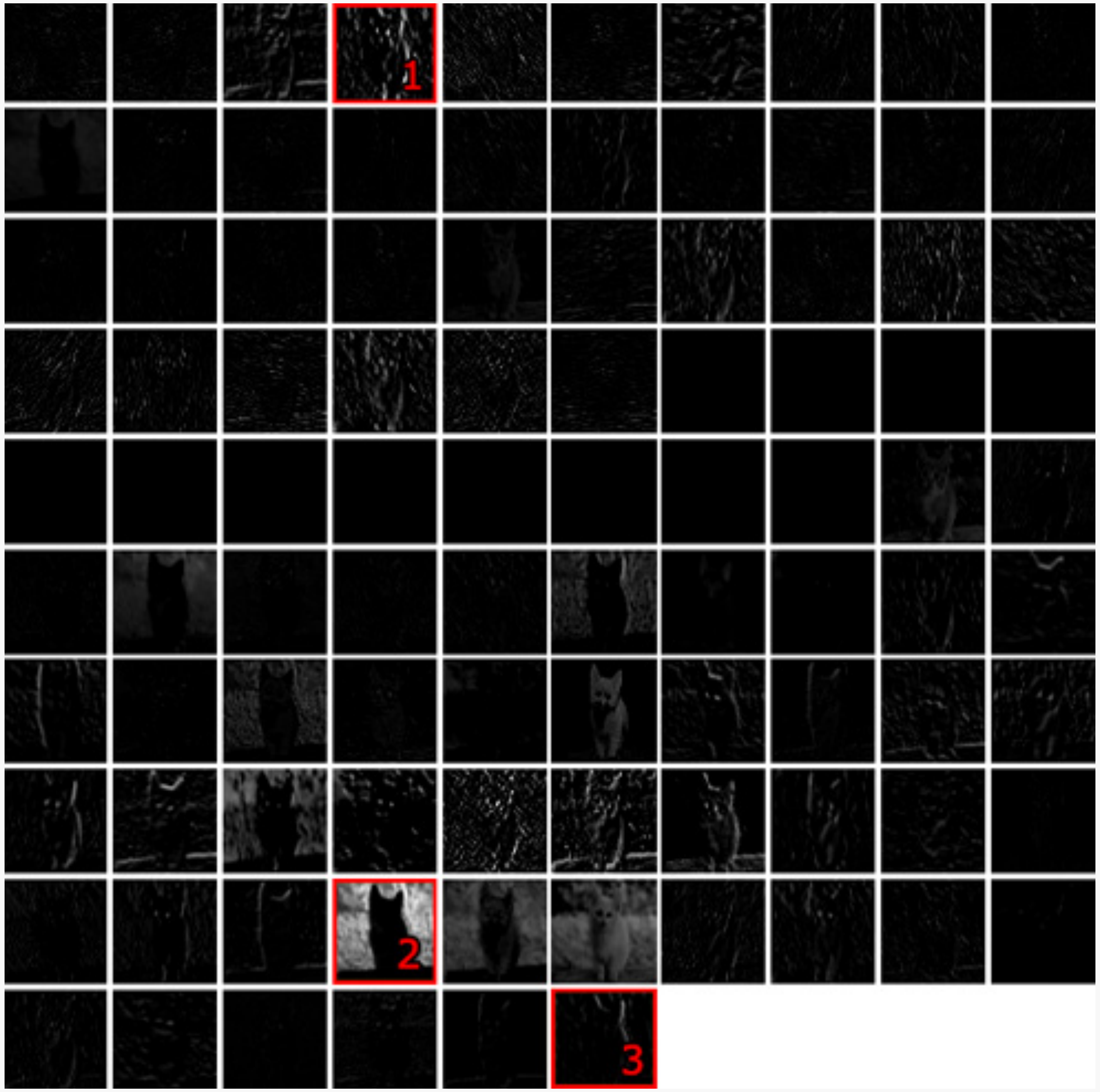
Deep Learning / Image Classification



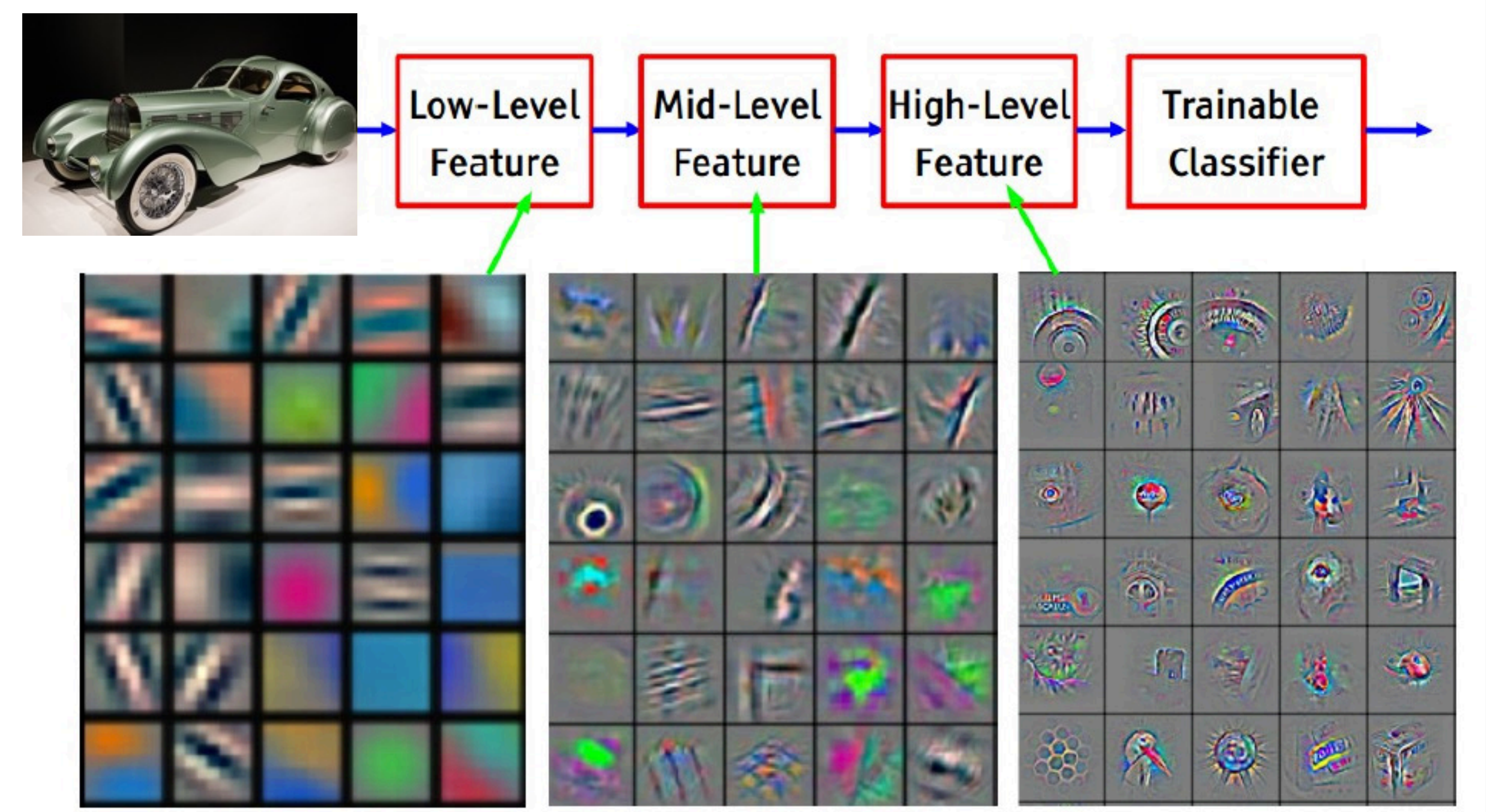


Filters and Activations



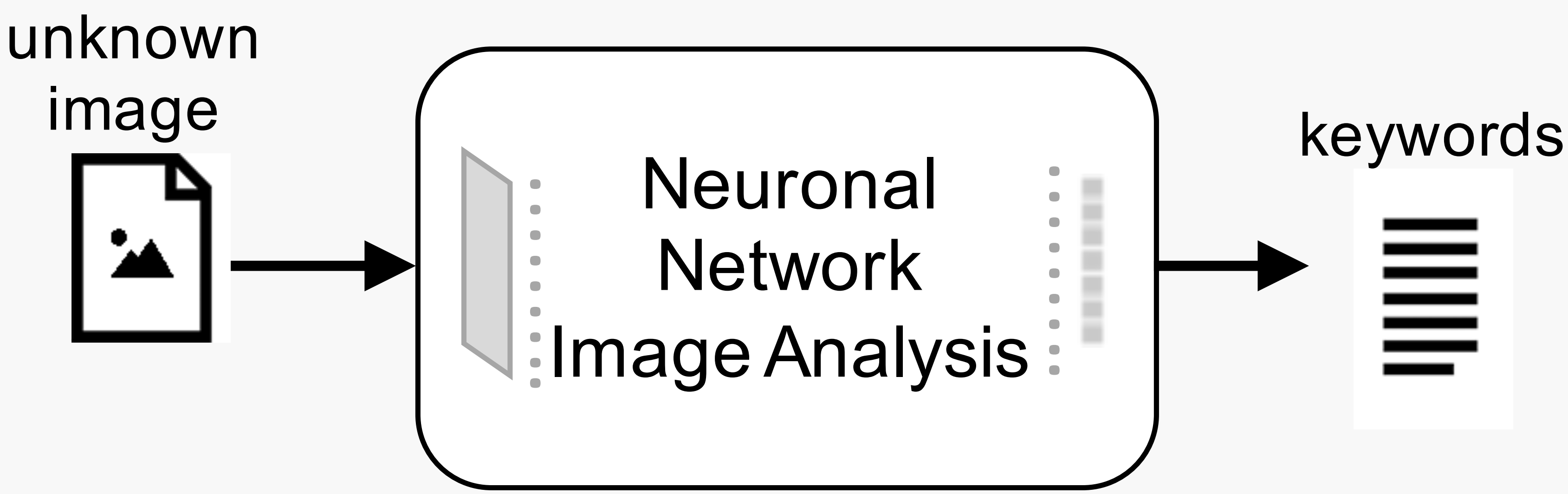


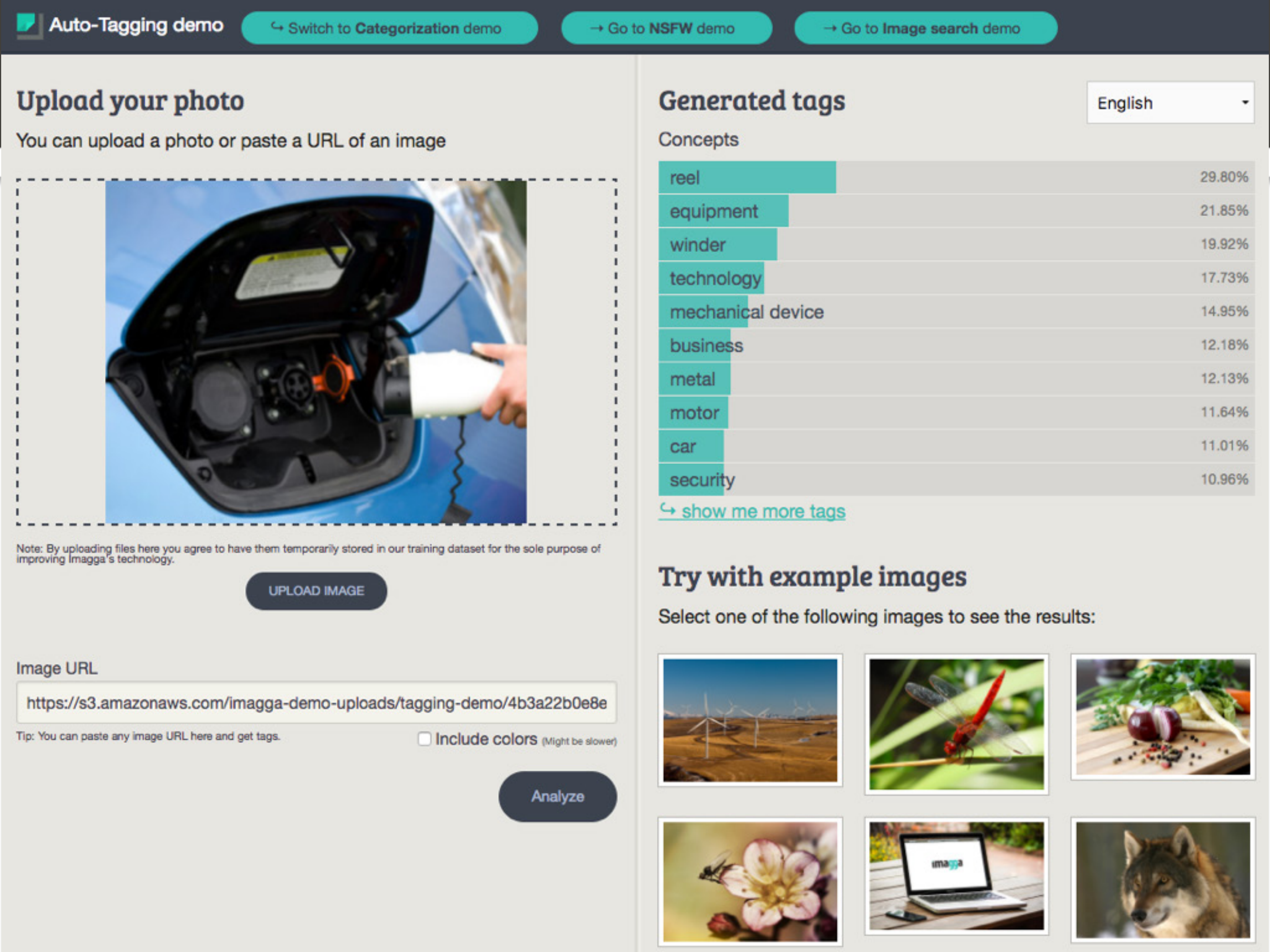
Convolutional Neural Networks



Automatic Image Tagging

Typical ML-based keywording





Auto-Tagging demo
↪ Switch to Categorization demo
→ Go to NSFW demo
→ Go to Image search demo
Upload your photo
You can upload a photo or paste a URL of an image
Note: By uploading files here you agree to have them temporarily stored in our training dataset for the sole purpose of improving Imagga's technology.
UPLOAD IMAGE
Image URL
https://s3.amazonaws.com/imagga-demo-uploads/tagging-demo/4b3a22b0e8e
Tip: You can paste any image URL here and get tags.
Include colors (Might be slower)
Analyze
Generated tags
English
Concepts
reel 29.80%
equipment 21.85%
winder 19.92%
technology 17.73%
mechanical device 14.95%
business 12.18%
metal 12.13%
motor 11.64%
car 11.01%
security 10.96%
↪ show me more tags
Try with example images
Select one of the following images to see the results:

clarifai Demo Solutions Pricing Developer API Resources Get your free API key Talk to us

Clarifai Demo GENERAL FACE NSFW COLOR MORE MODELS

TRY YOUR OWN IMAGE OR VIDEO

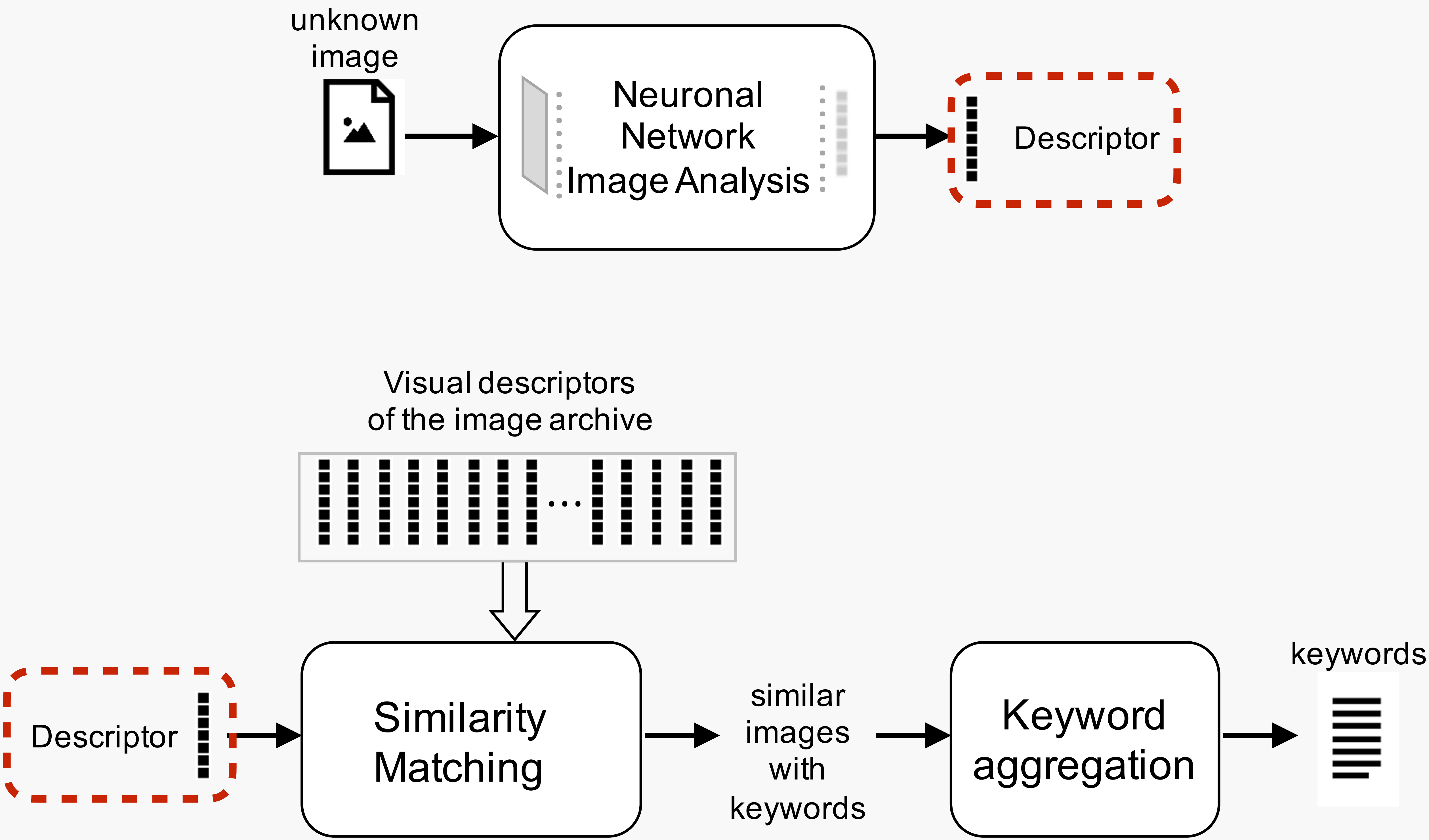
General VIEW DOCS

LANGUAGE

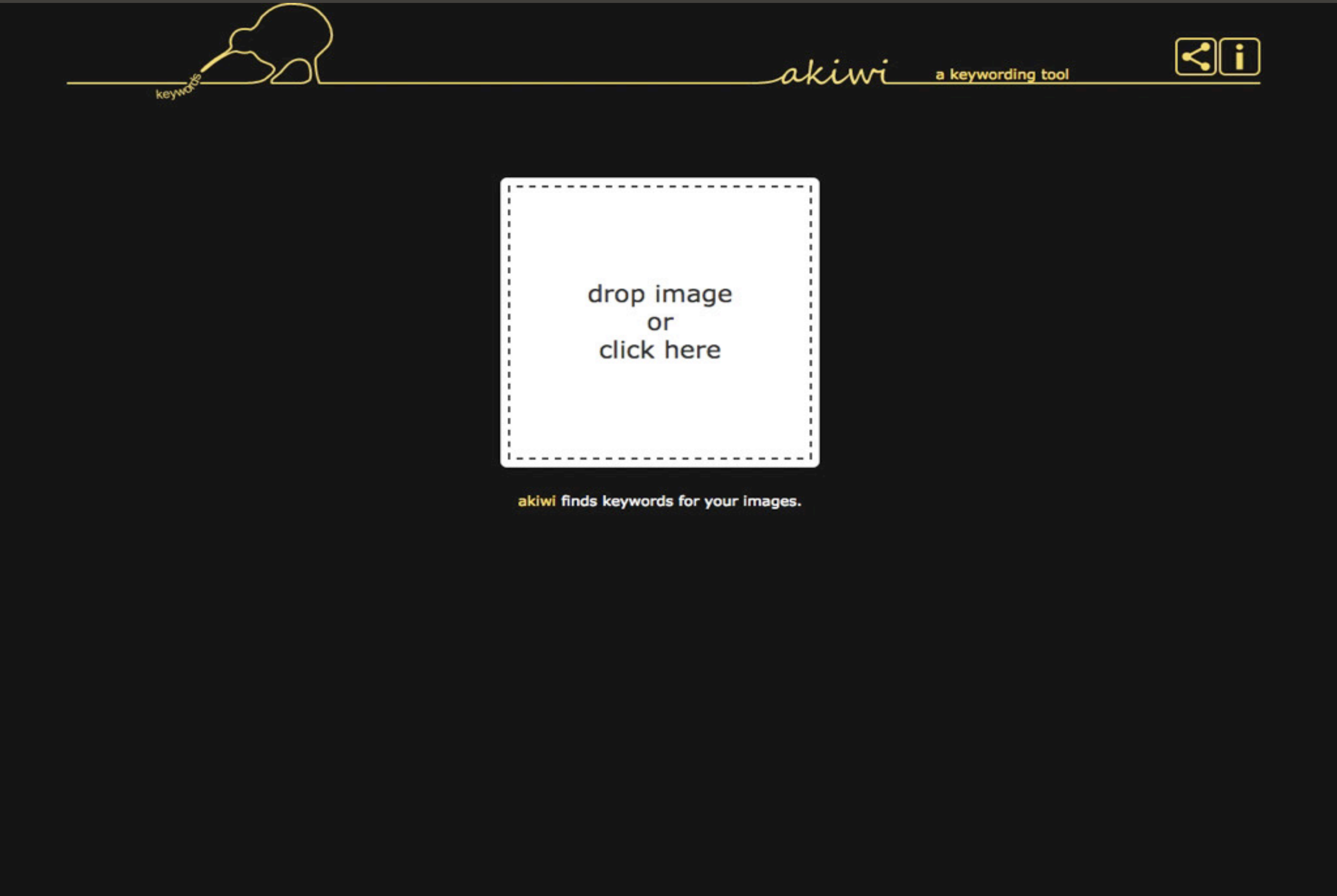
English (en)

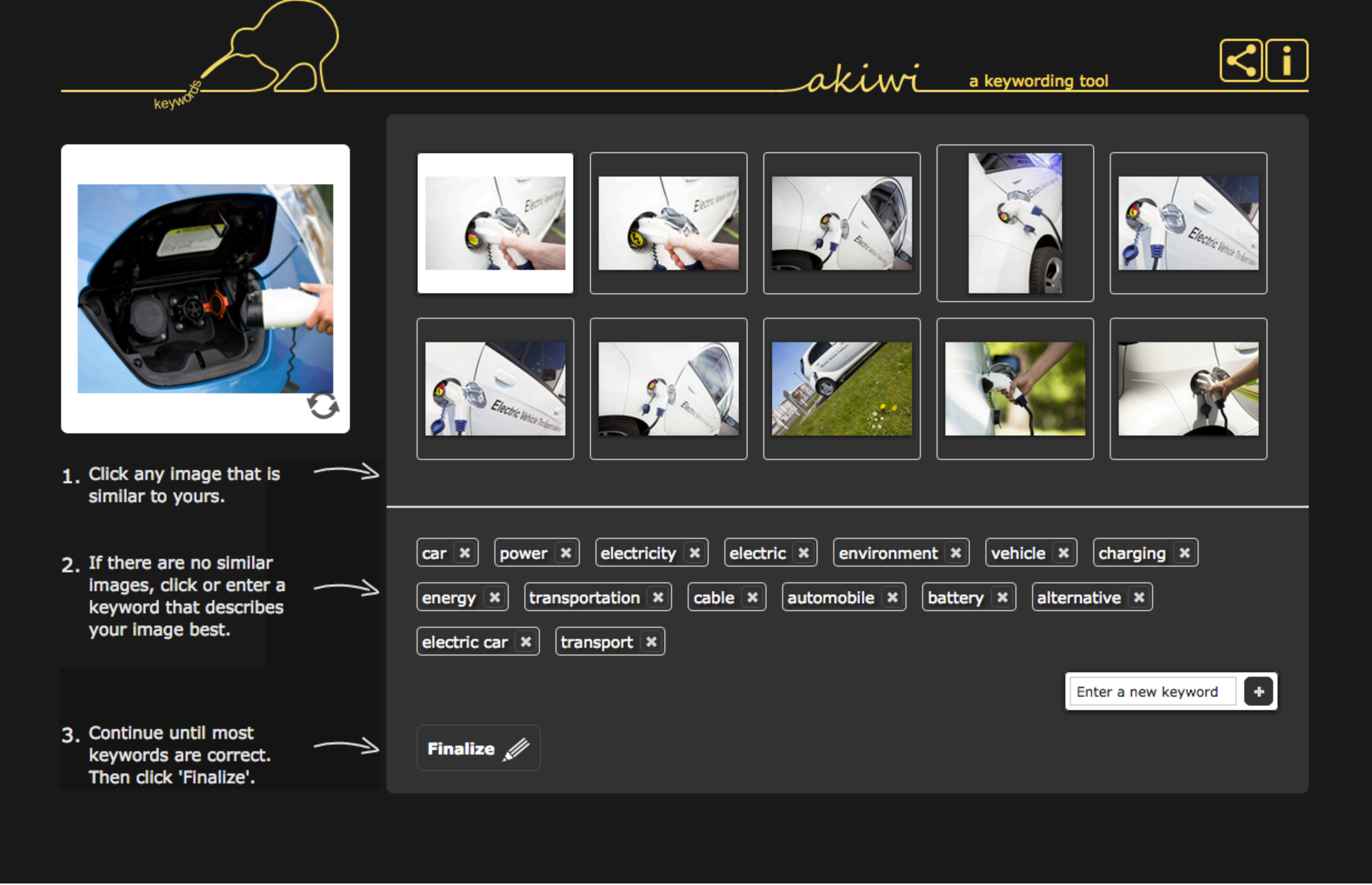
PREDICTED CONCEPT	PROBABILITY
car	0.999
vehicle	0.997
drive	0.996
transportation system	0.995
power	0.985
classic	0.981
front	0.980
speed	0.976
technology	0.975
wheel	0.973
engine	0.970
old	0.940
fast	0.938
chrome	0.935
retro	0.928
steering wheel	0.927
cockpit	0.927
glazed	0.927

Use your own images & keywords



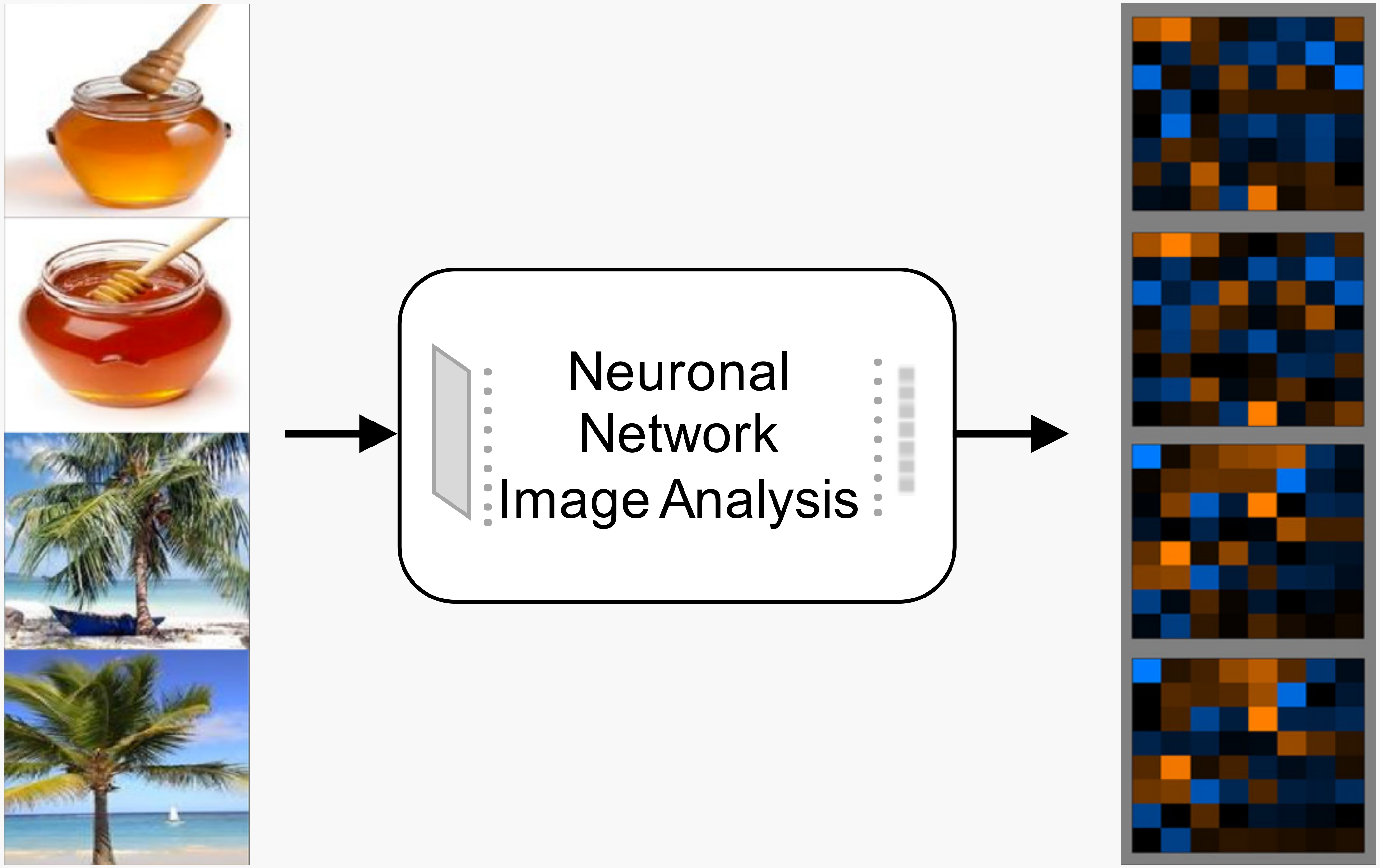
Semi-automatic Tagging akiwi.eu




keywords
akiwi
a keywording tool
1. Click any image that is similar to yours.
2. If there are no similar images, click or enter a keyword that describes your image best.
3. Continue until most keywords are correct. Then click 'Finalize'.
car
power
electricity
electric
environment
vehicle
charging
energy
transportation
cable
automobile
battery
alternative
electric car
transport
Enter a new keyword
Finalize

New Image Descriptors

Image descriptors from NN



What are Neural Networks good at?

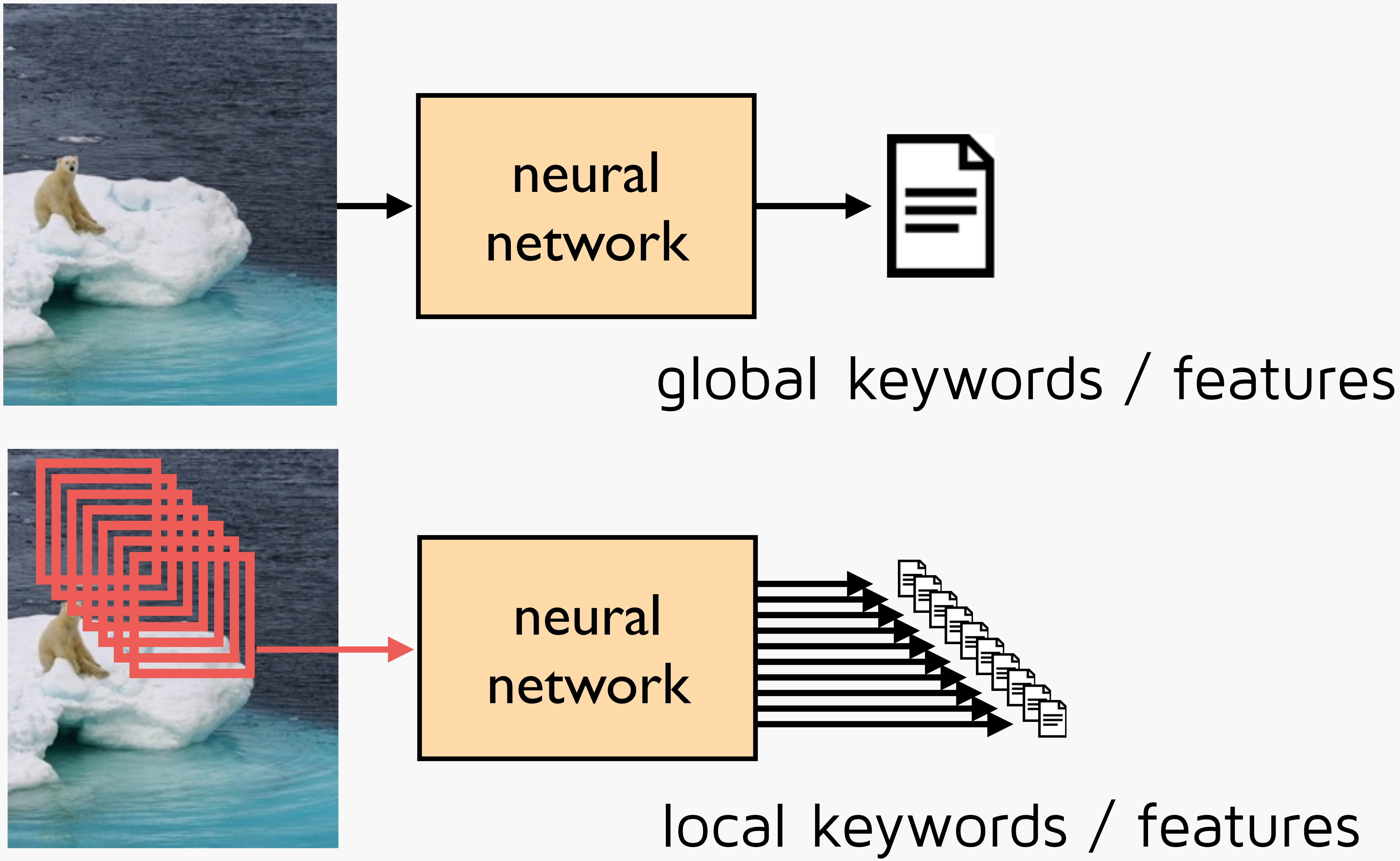
Recognition of

- Objects (type, manufacturer, ...)
- Humans (sex, age, mood, posture, ...)
- Style, Artist, Composition, ...

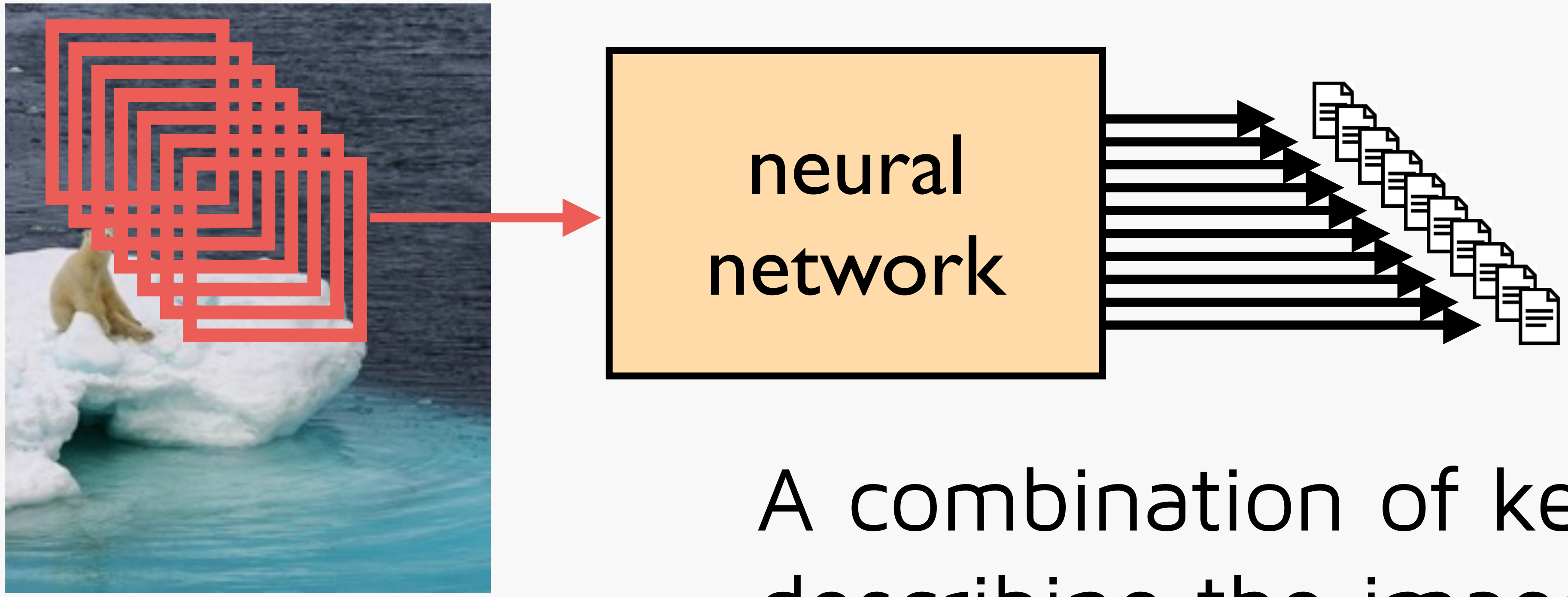
Everything a human (expert) could recognize in a few seconds.

If you have enough data to train the network

Global vs. local analysis



What kind of descriptors will we have?



A combination of keywords describing the image?

or

A collection of features characterizing image content, style, mood & look? (language independent)

Automatic Image Description

Long-term Recurrent Convolutional Networks for Visual Recognition and Description

Jeff Donahue* Lisa Anne Hendricks* Sergio Guadarrama* Marcus Rohrbach**
Subhashini Venugopalan† Kate Saenko‡ Trevor Darrell**

A female tennis player in action on the court.

A group of young men playing a game of soccer.

A man riding a wave on top of a surfboard.

Visual Image Browsing

Searching by scrolling

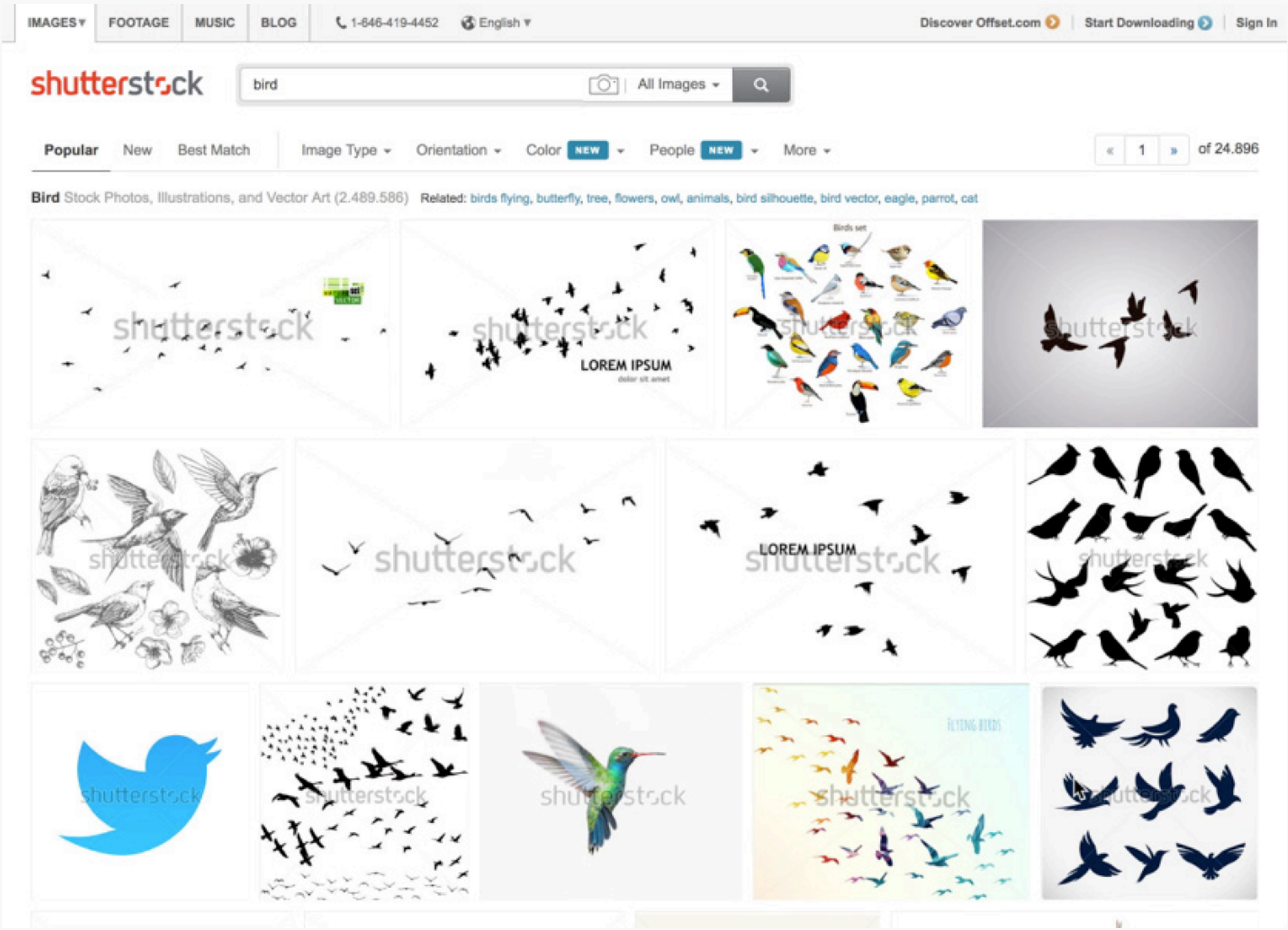
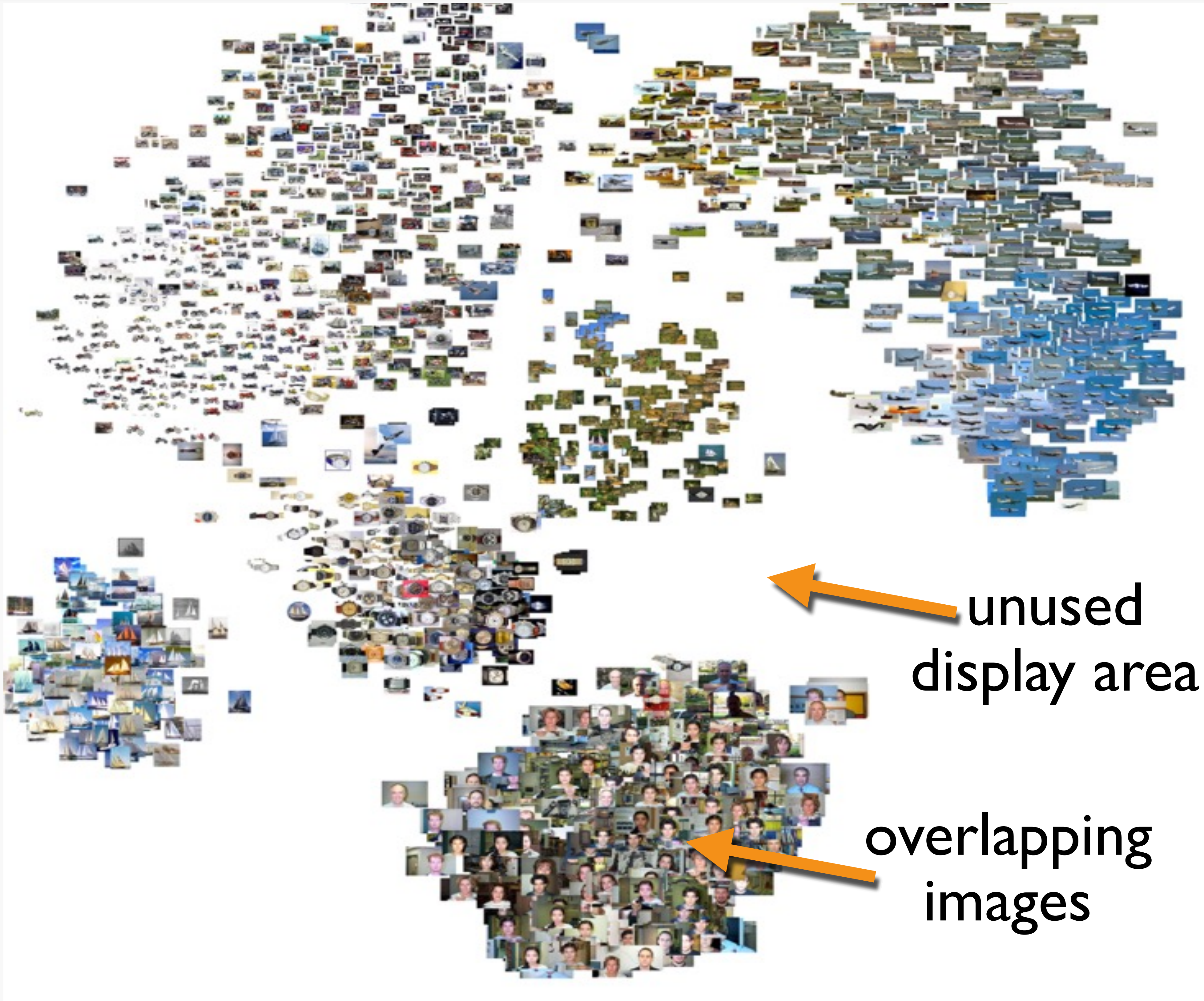
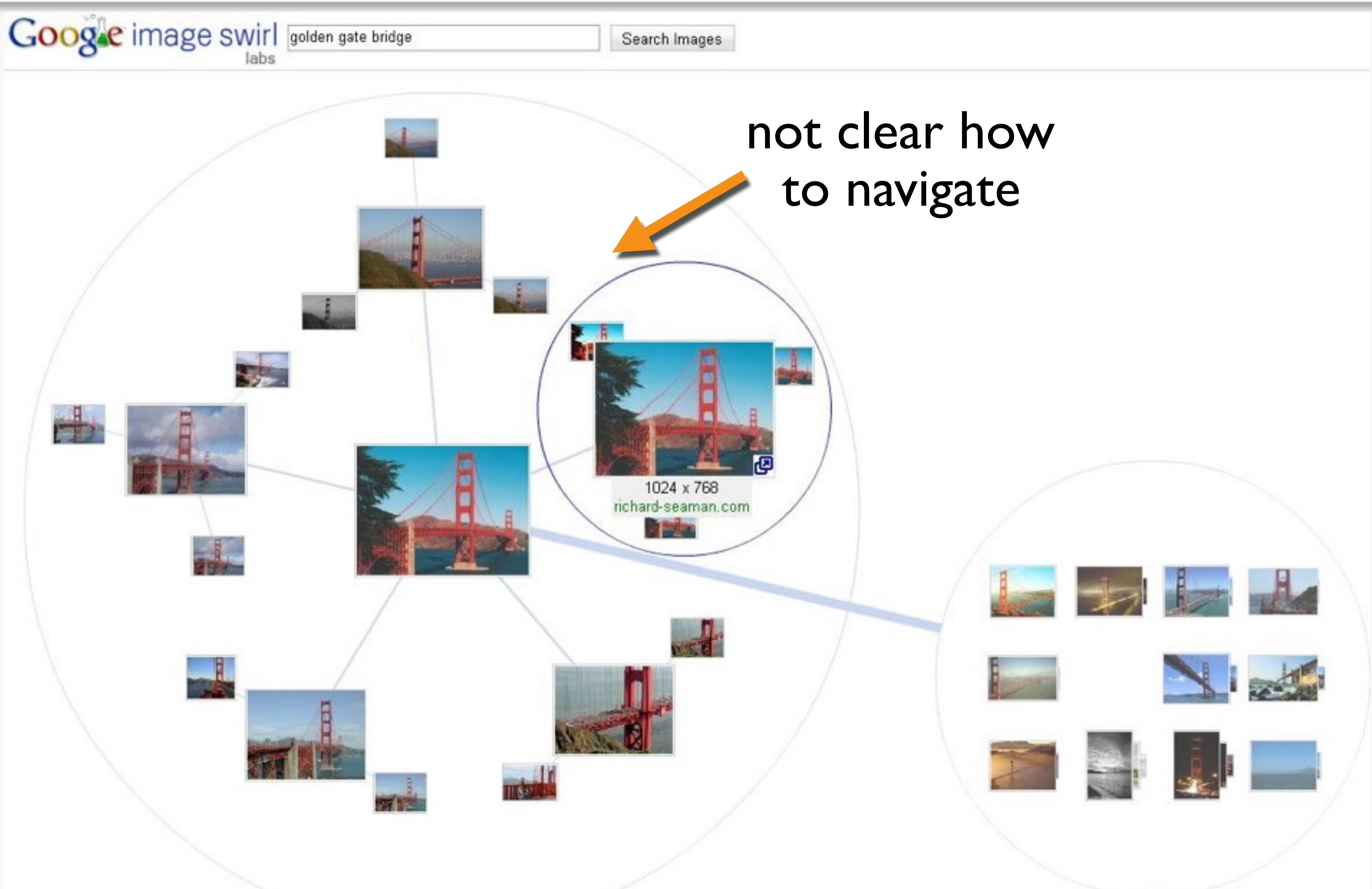


Image Projection using t-SNE



Google Image Swirl



not clear how to navigate

Requirements for image browsing

- Image layout/projection:
  - use entire display area, no overlap
  - good mapping (similar neighbor images)
- Image sets:
  - support for millions of (untagged) images
  - allow changes (deleted or added images)
- Navigation:
  - easy & natural
  - purely visual (without keywords)
  - hierarchical approach

Our initial approach

based on the idea of mapping services

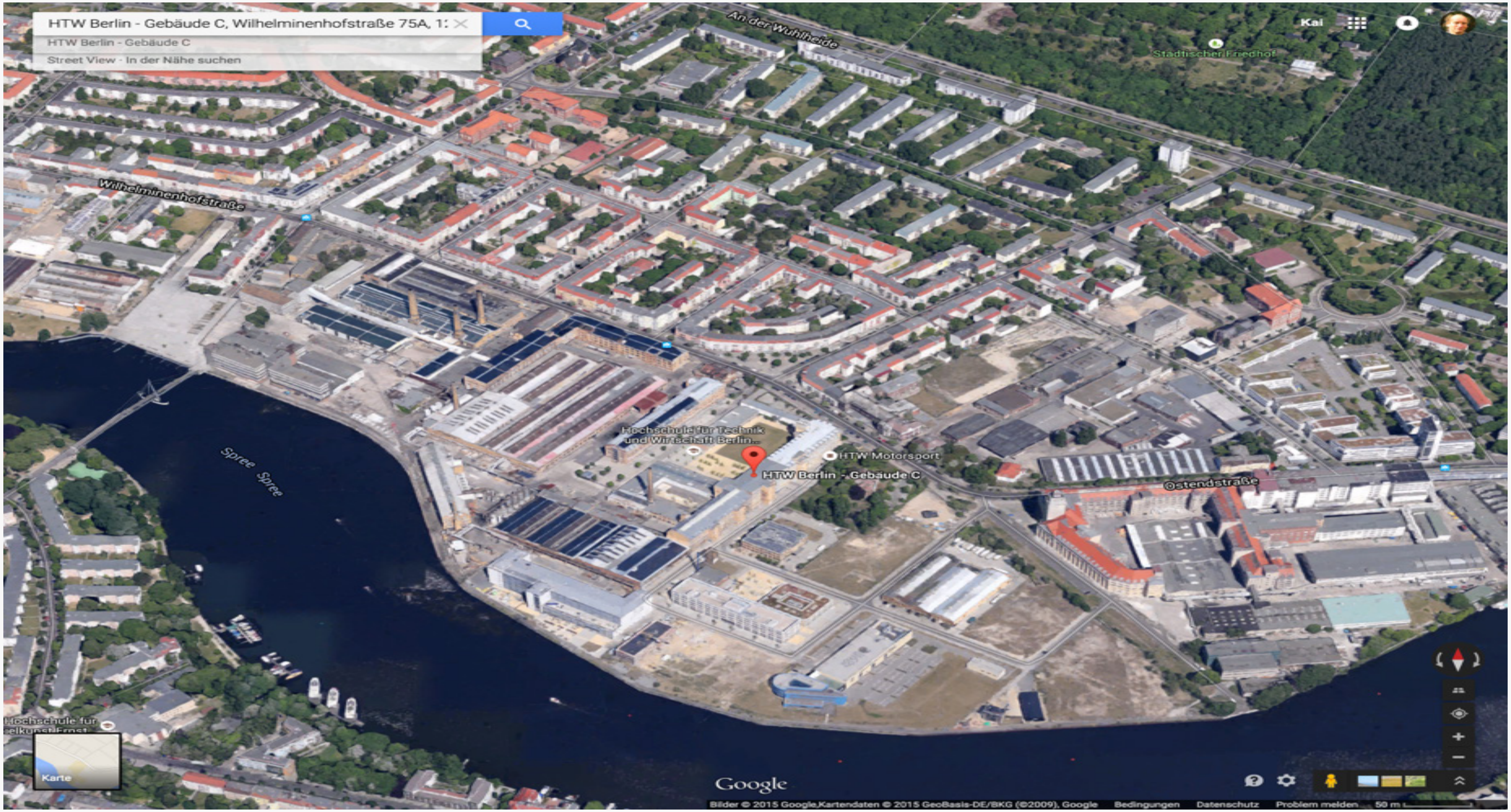
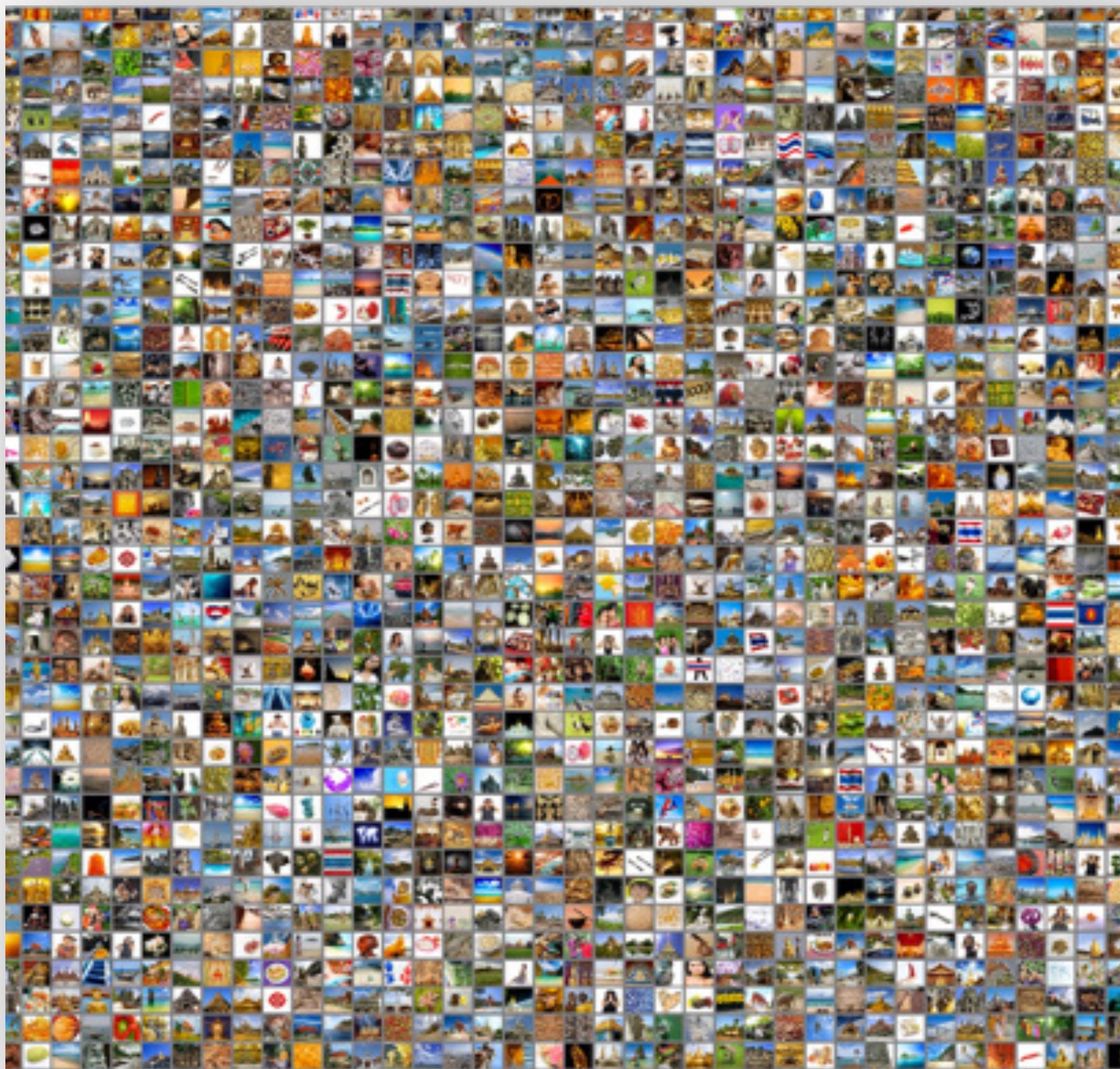
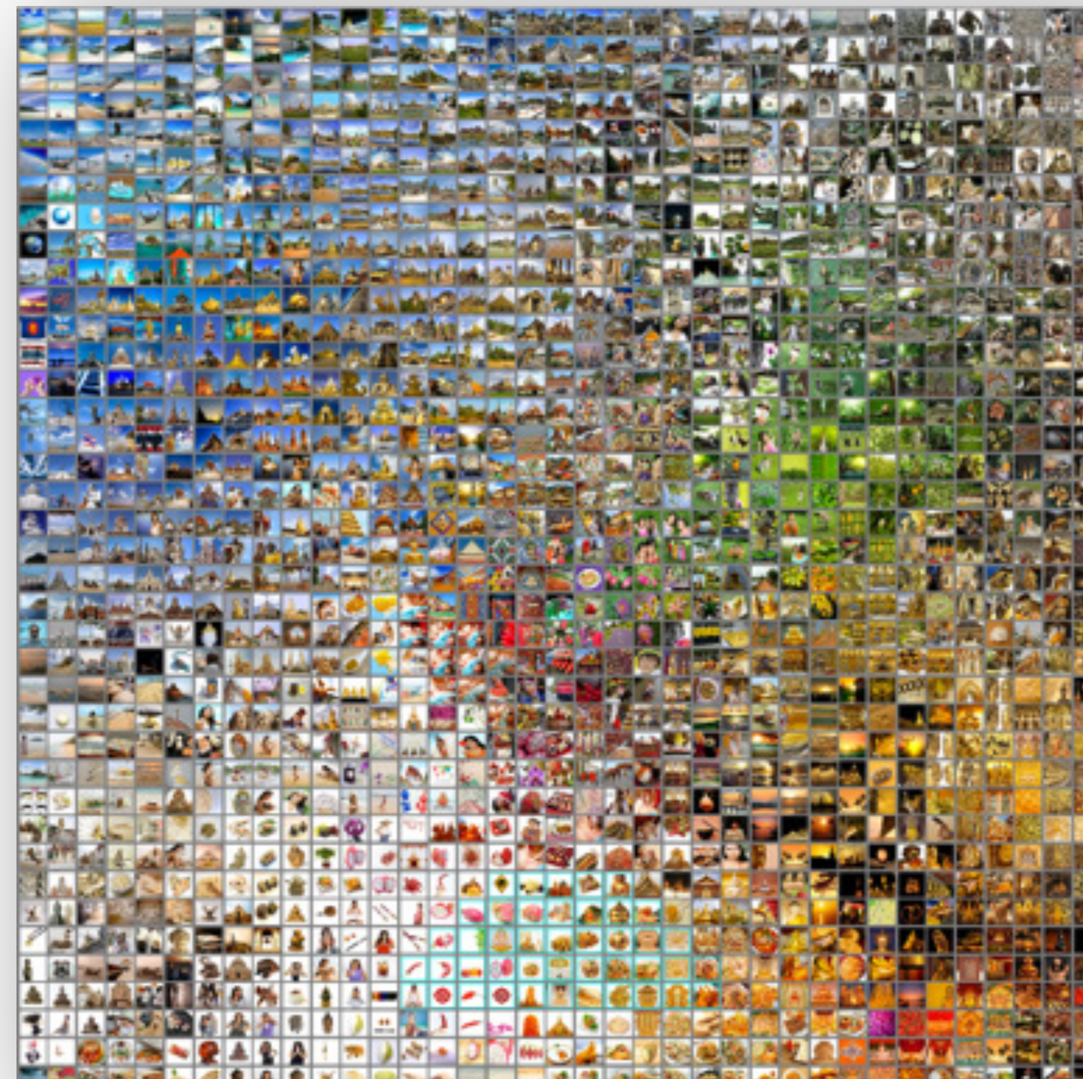


Image sorting using a fast hierarchical SOM

unsorted

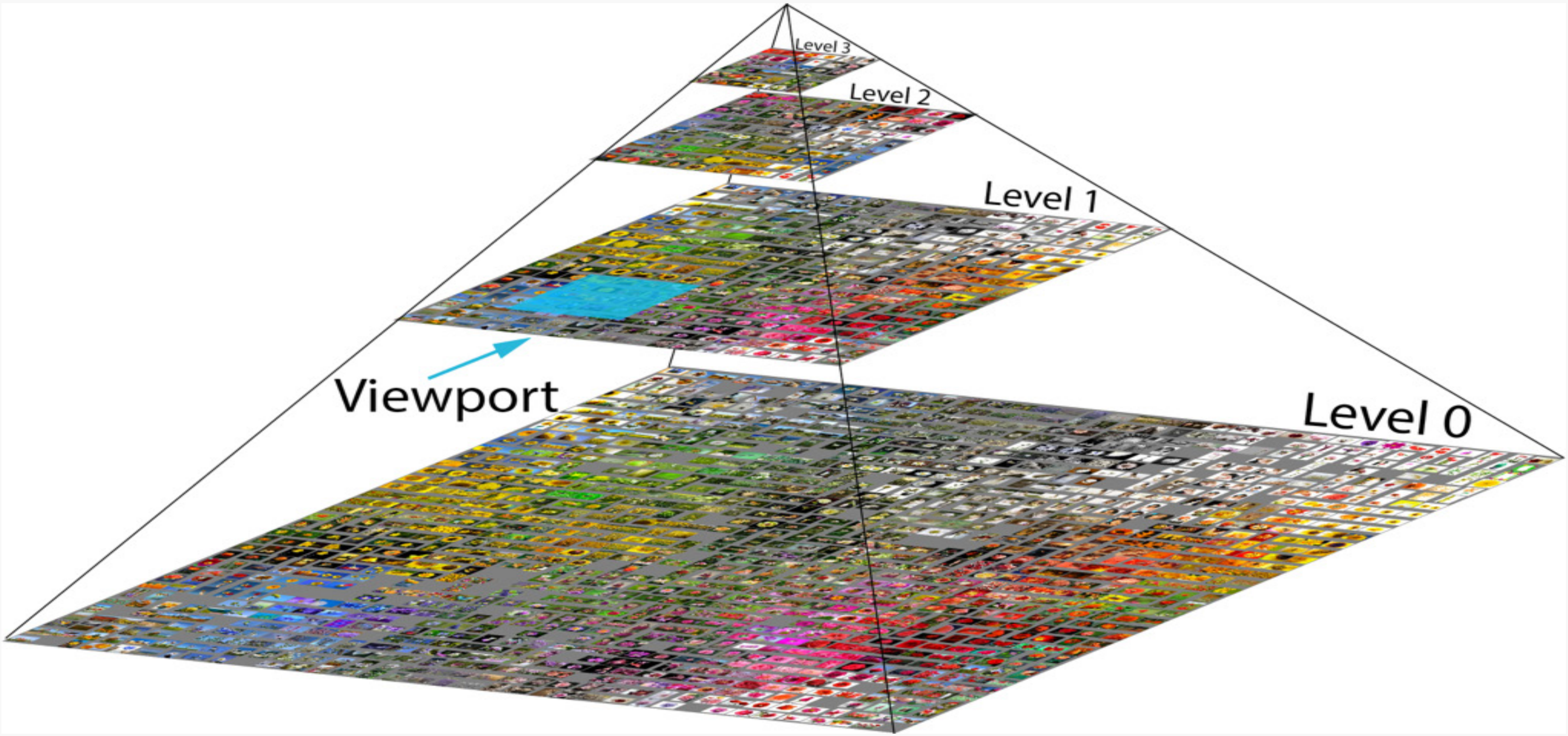


sorted

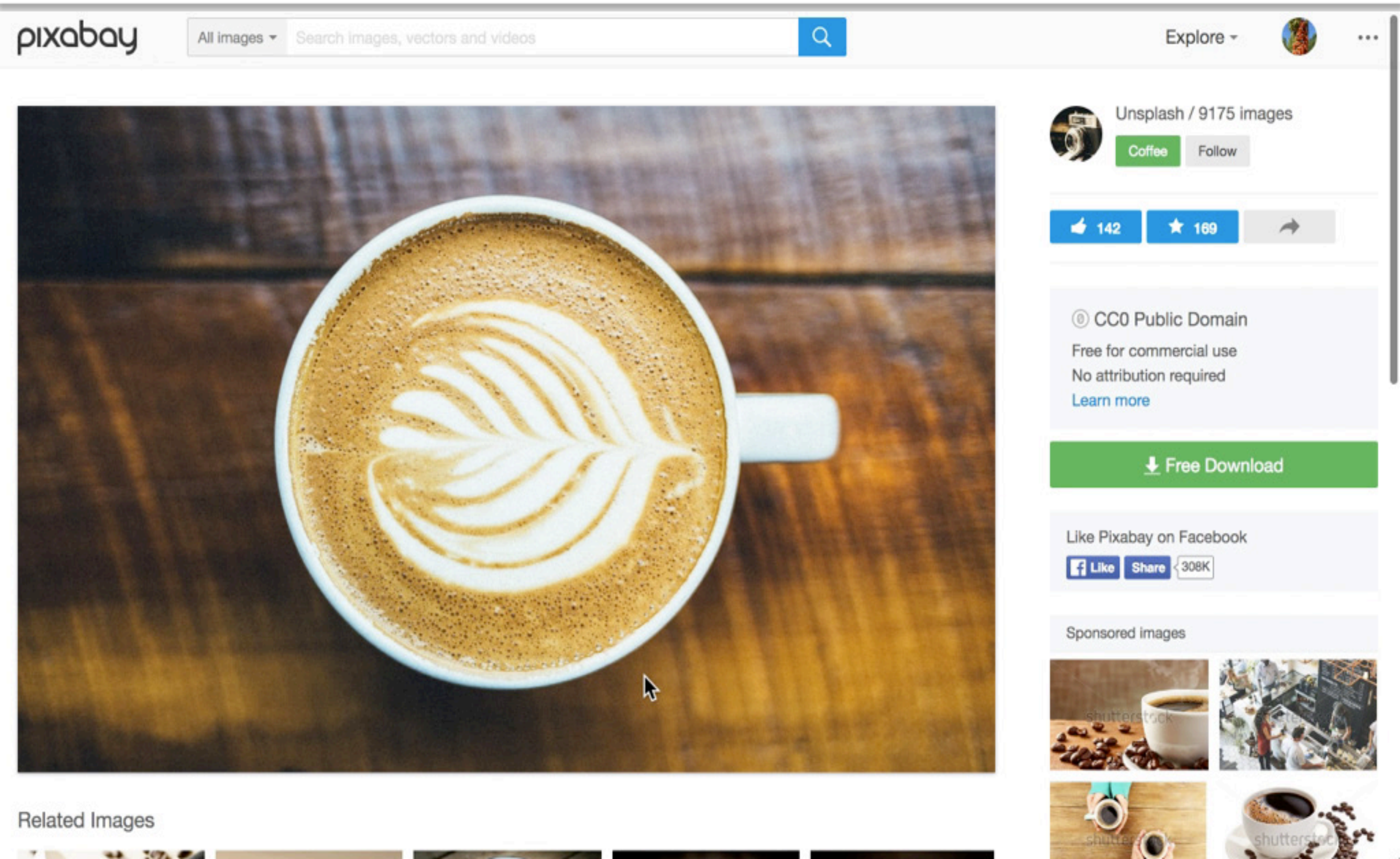


complexity $O(n\ log(n))$, 50 times faster than Barnes-Hut t-SNE, similar projection quality

Hierarchical image layers



picsbuffet.com @ pixabay



Evaluation of picsbuffet

+ Good visualization & easy navigation
+ Very fast (no server calculations)
+ Suited for very large image sets

– Image relationships are too complex to be mapped to 2D
– Bad for changing image sets

→ image graphs

– Sometimes poor neighbors

→ improved image features

Requirements for improved image browsing

1. High quality semantic image features to model image similarities well
2. Fast construction of “good“ hierarchical image graphs
3. Good graph visualization and easy to use navigation techniques

2. Graph requirements

A. Connected images should be similar.
B. The number of connections per image should not be too high or too low.
C. The path between any two images should be as short as possible.

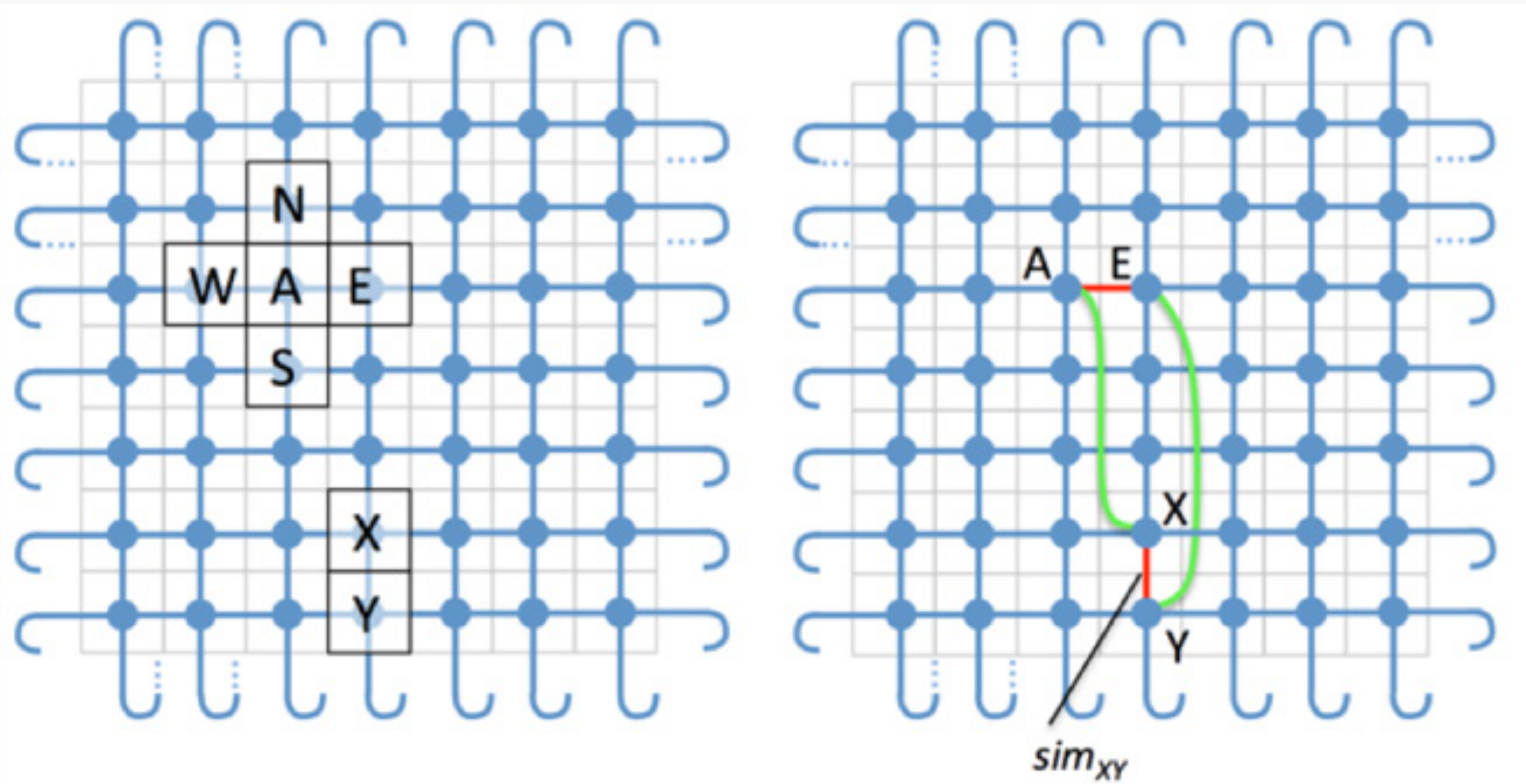
Very hard optimization problem!

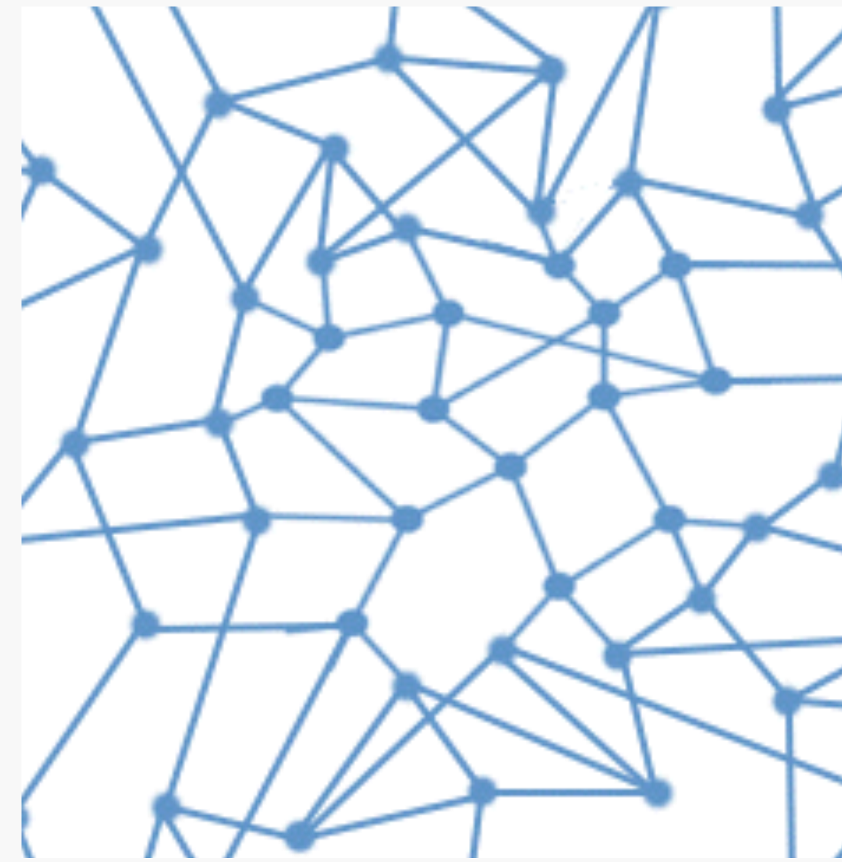
➡ A: ensured by
good image features & graph building

➡ B: fixed to 4 connections

➡ C: simplified by checking the connectivity

2. Building the graph



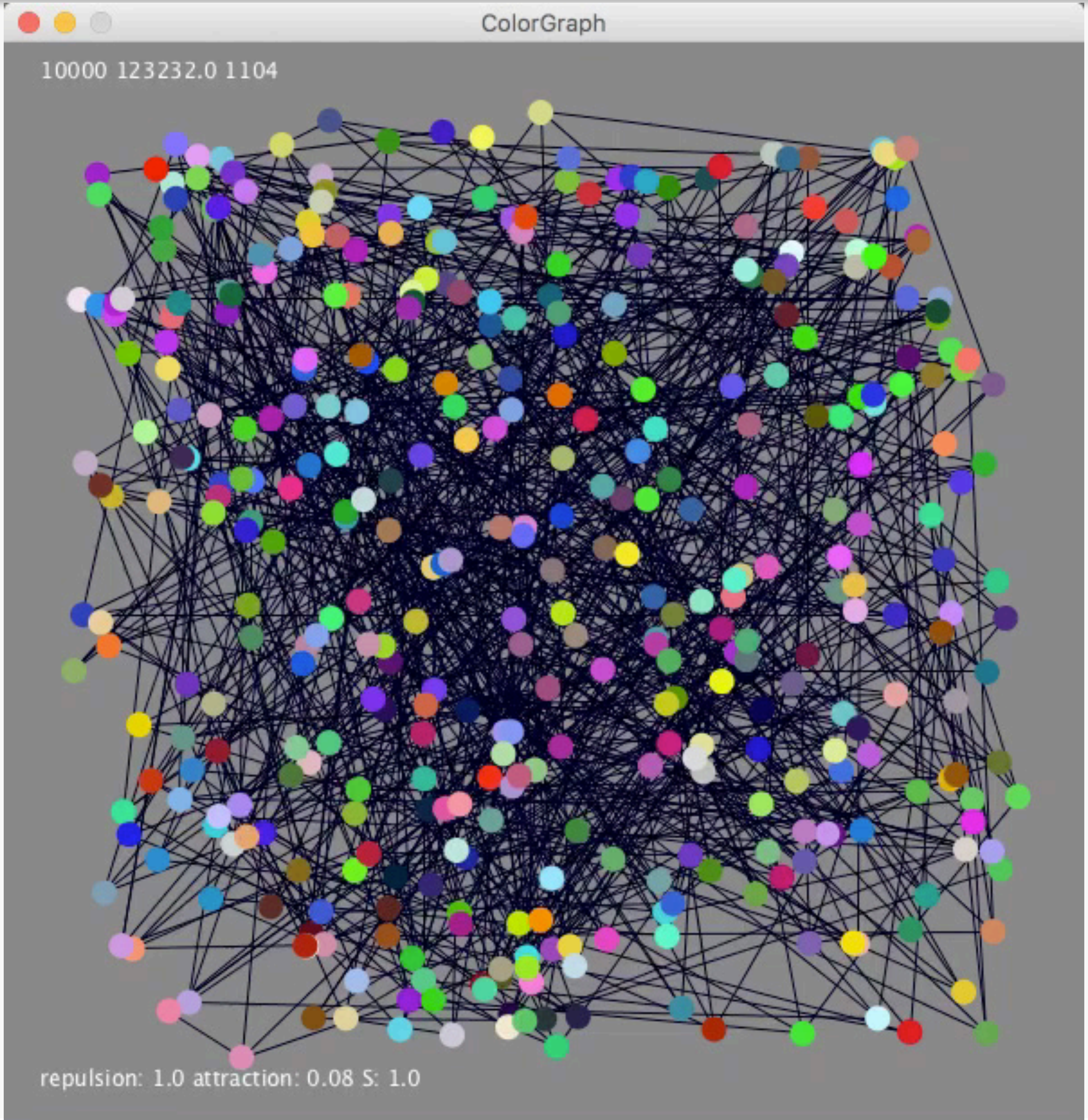


Starting from a regular mesh
we swap edges if

$$sim_{AX} + sim_{EY} - sim_{AE} - sim_{XY} > 0$$

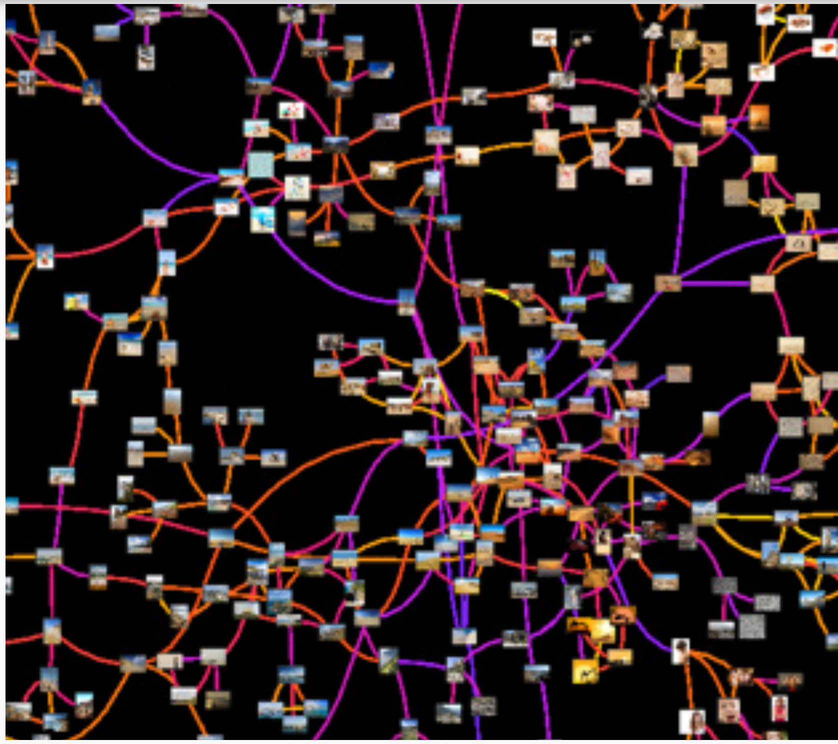
Finds good graphs, but convergence is slow

Better swapping, RGB color example



3. Graph display & navigation

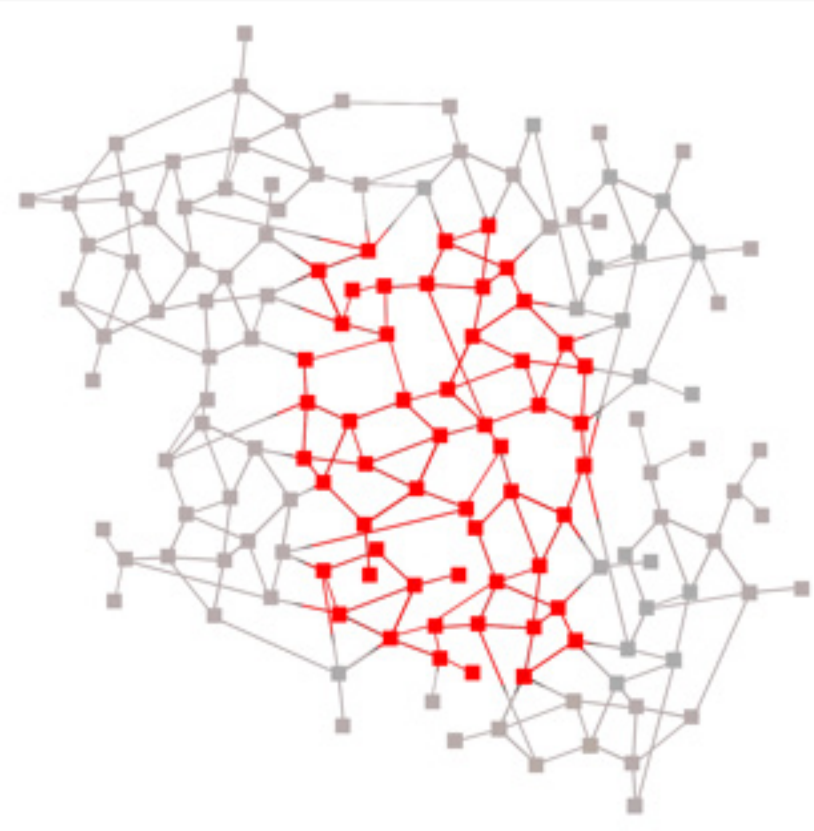
Visualizing and navigating graphs is difficult.
Layout algorithms are slow.
Radial views are confusing.

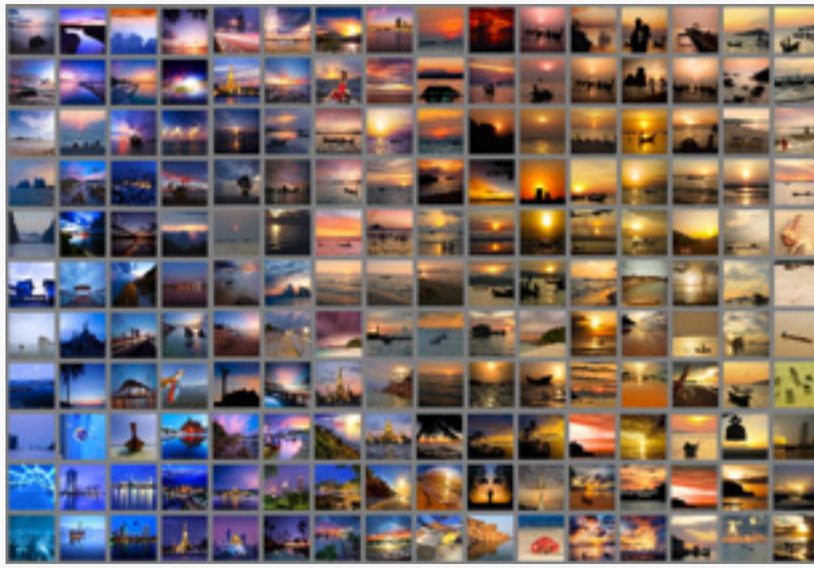


3. Our approach for graph navigation

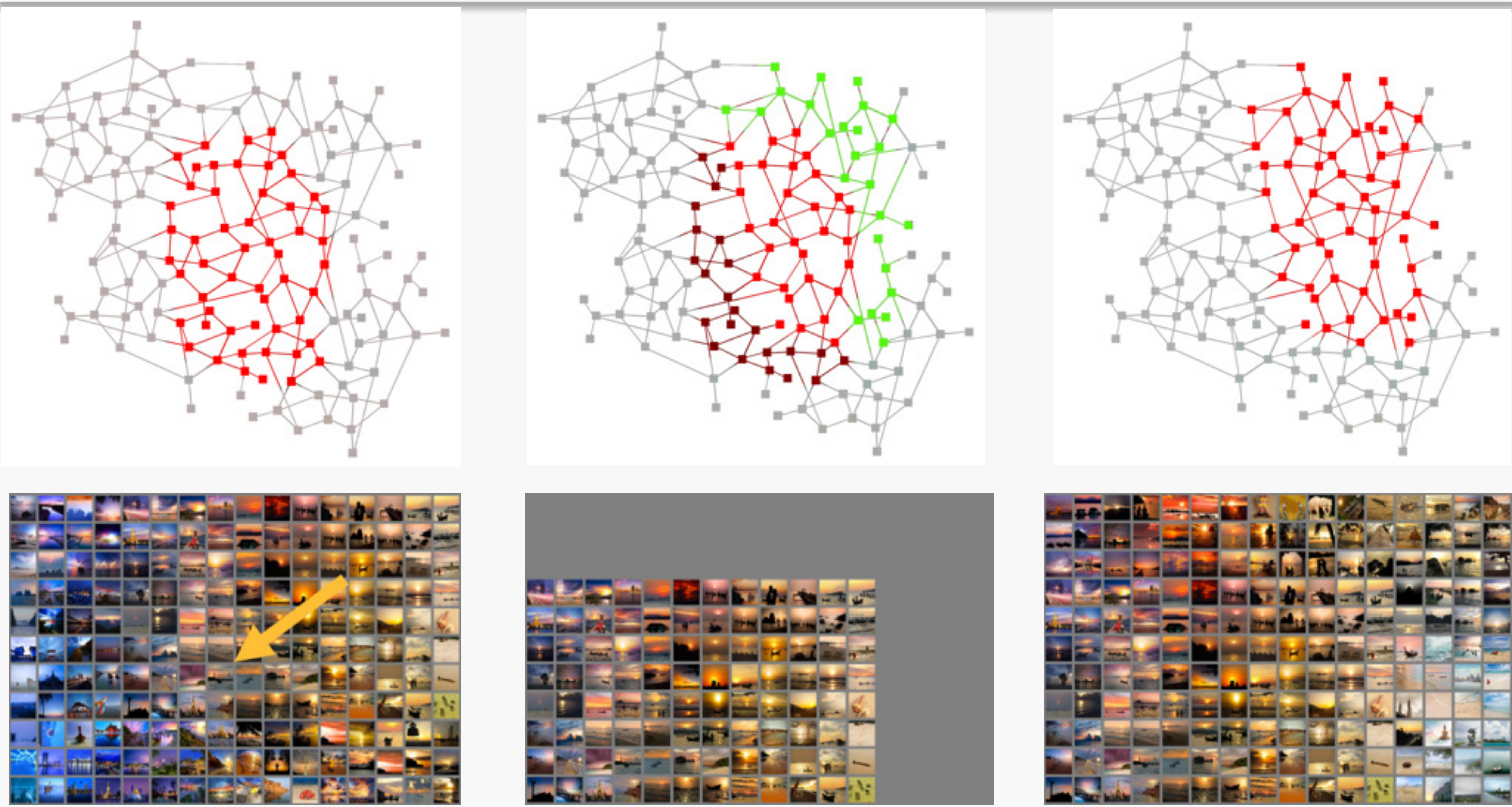
- Project fractions of the graph to a 2D map and perform instant 2D sorting

- Zooming is performed by blending to a graph of the next level



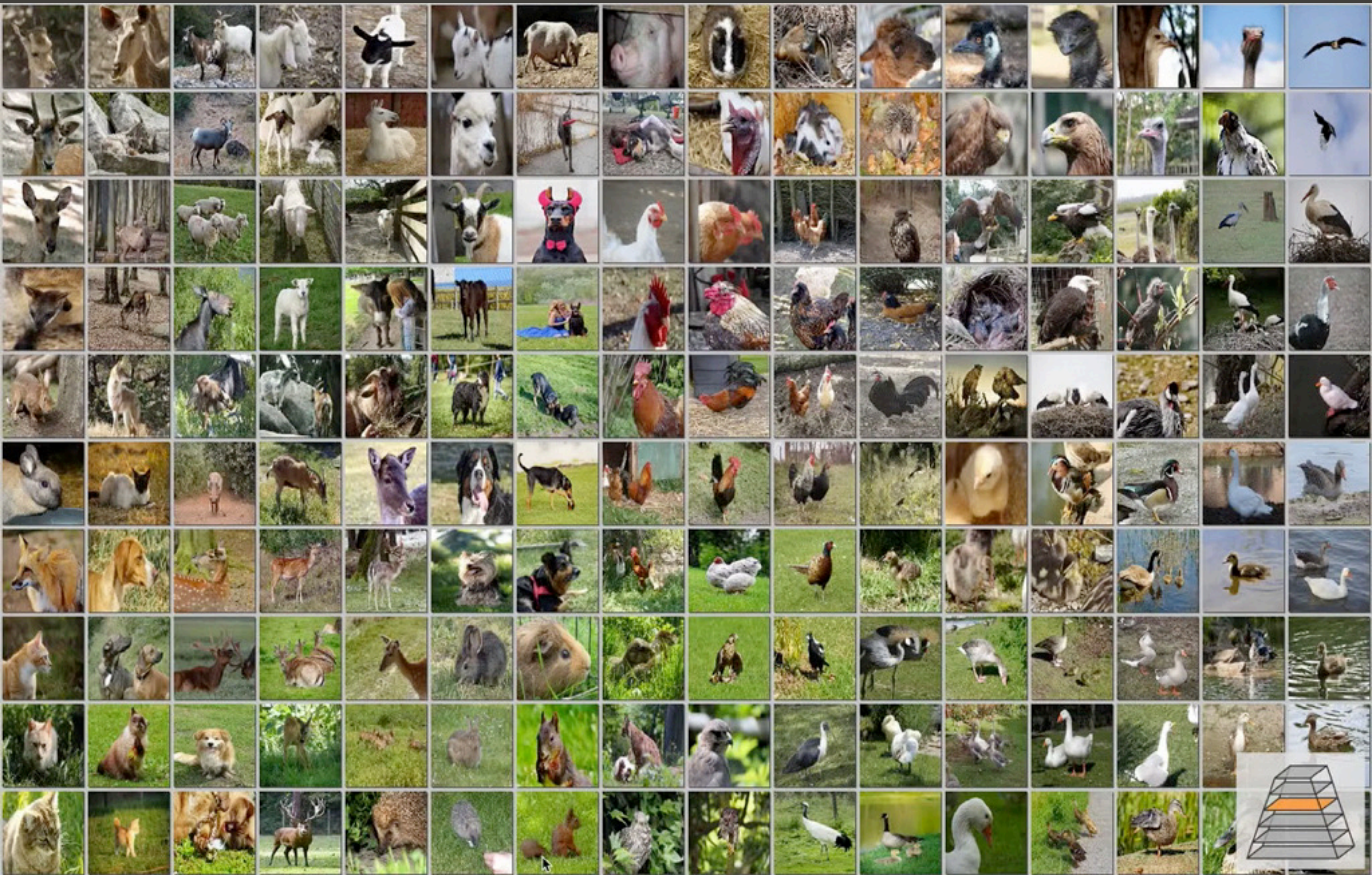


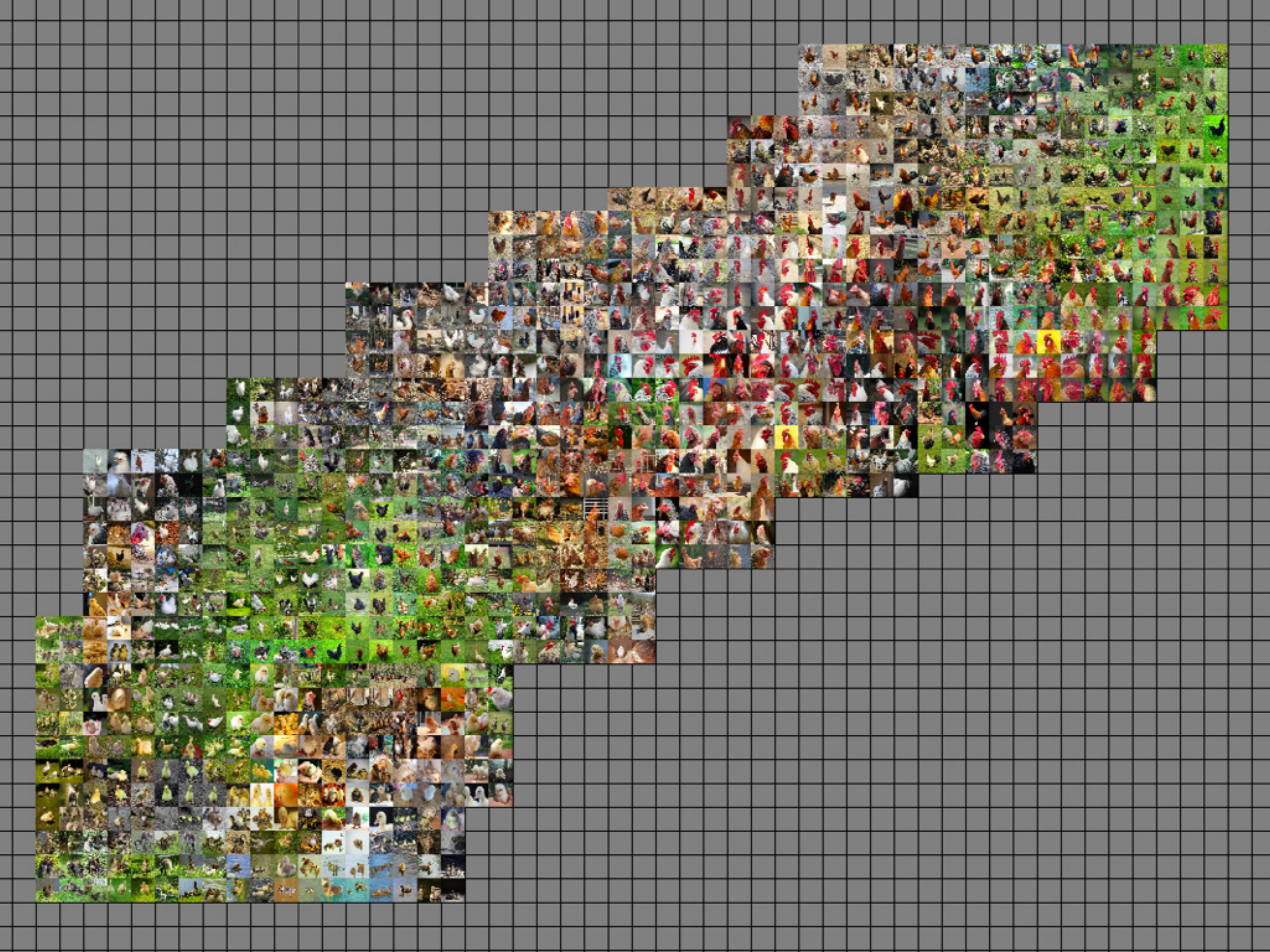
3. Graph Navigation



Moving the map retrieves new images from the graph.

Demo of Graph-based Browsing





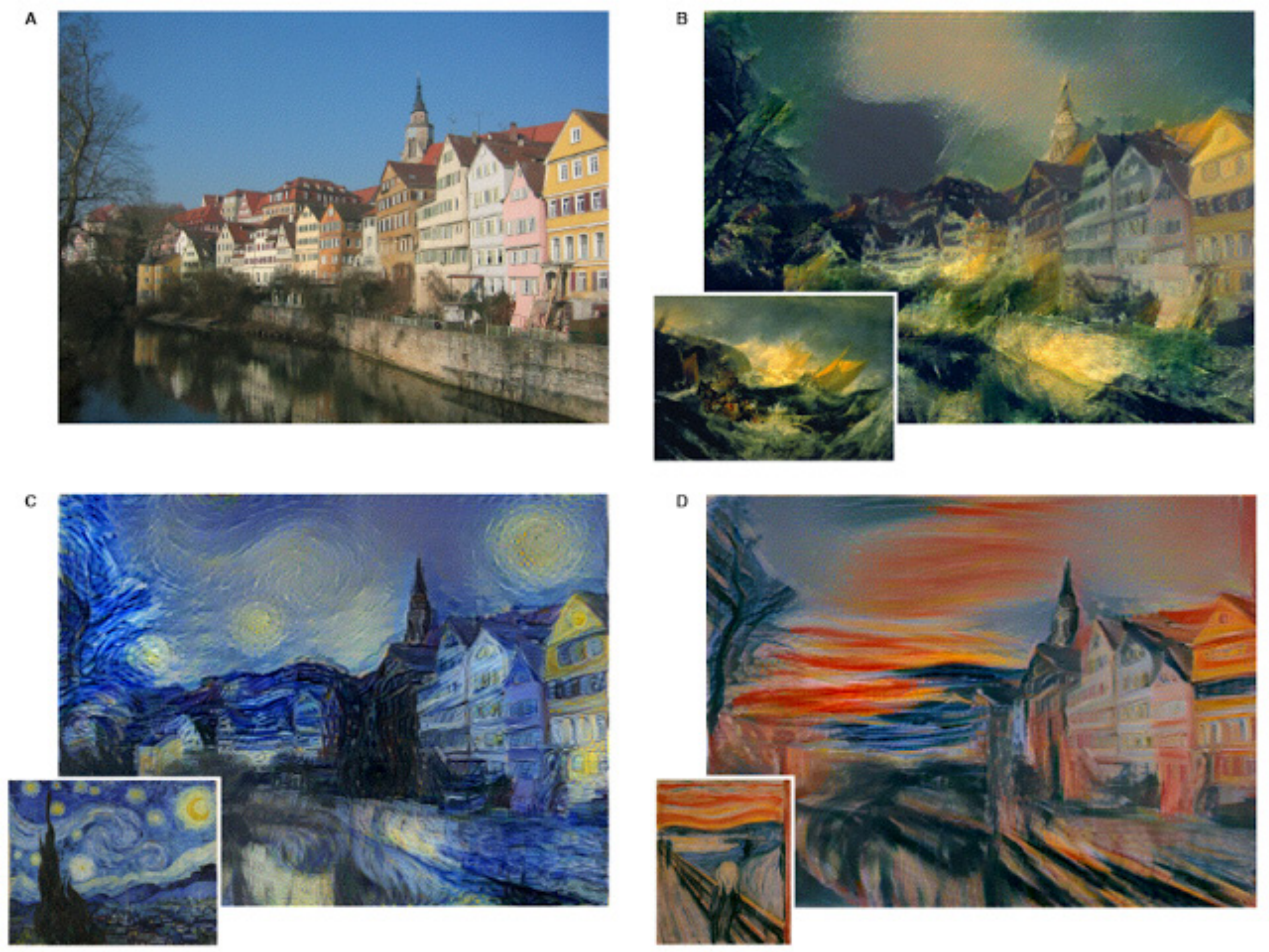
New Image Manipulation Tools
&
Image Synthesis

If you do not have the image you want ...

- Render it from a 3D model
- Change an existing image
- Synthesize it

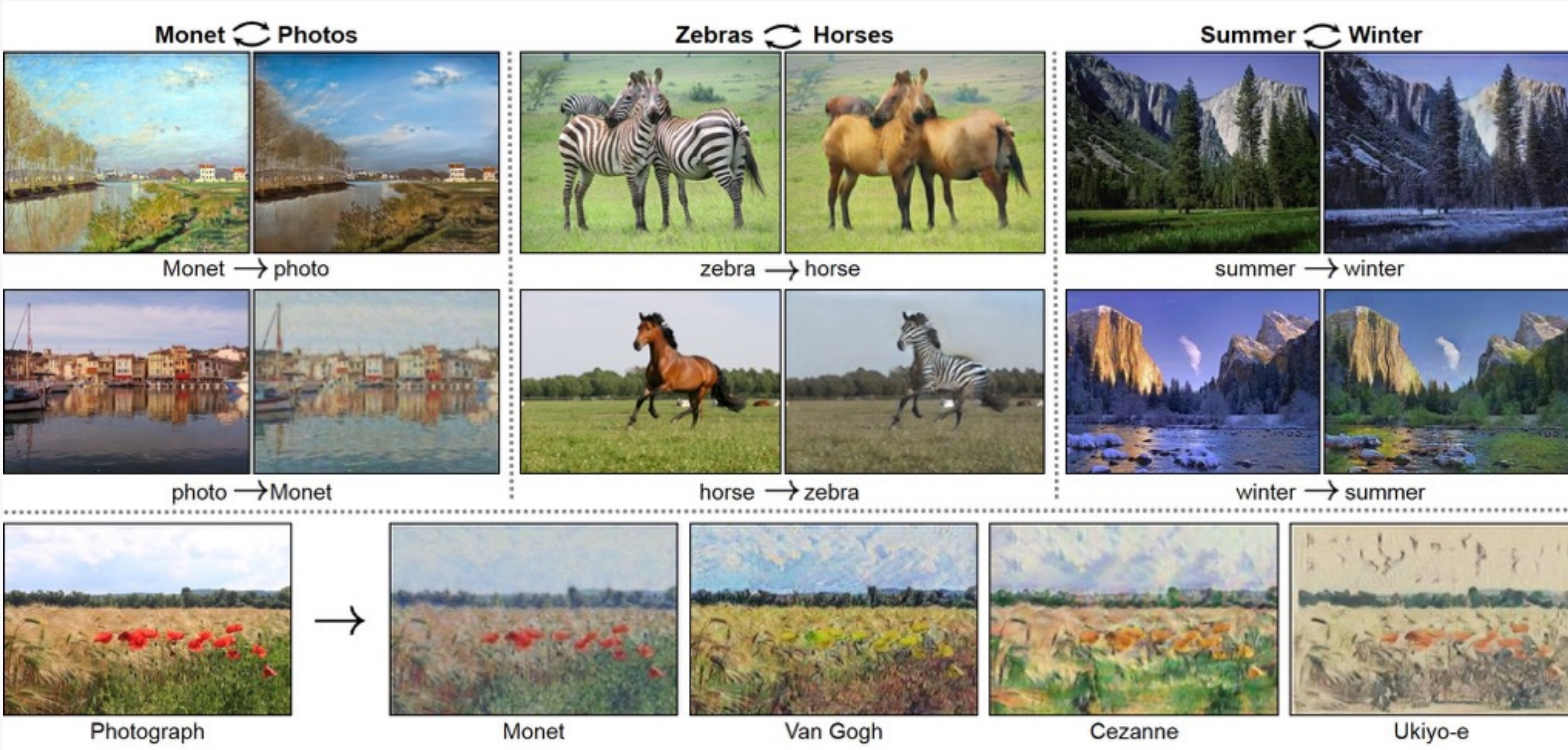
Rendered Images (IKEA)

Style Transfer deepart.io



Leon A. Gatys, Alexander S. Ecker, Matthias Bethge

Image Manipulation II



Jun-Yan Zhu   Taesung Park   Phillip Isola   Alexei A. Efros

Image Manipulation III

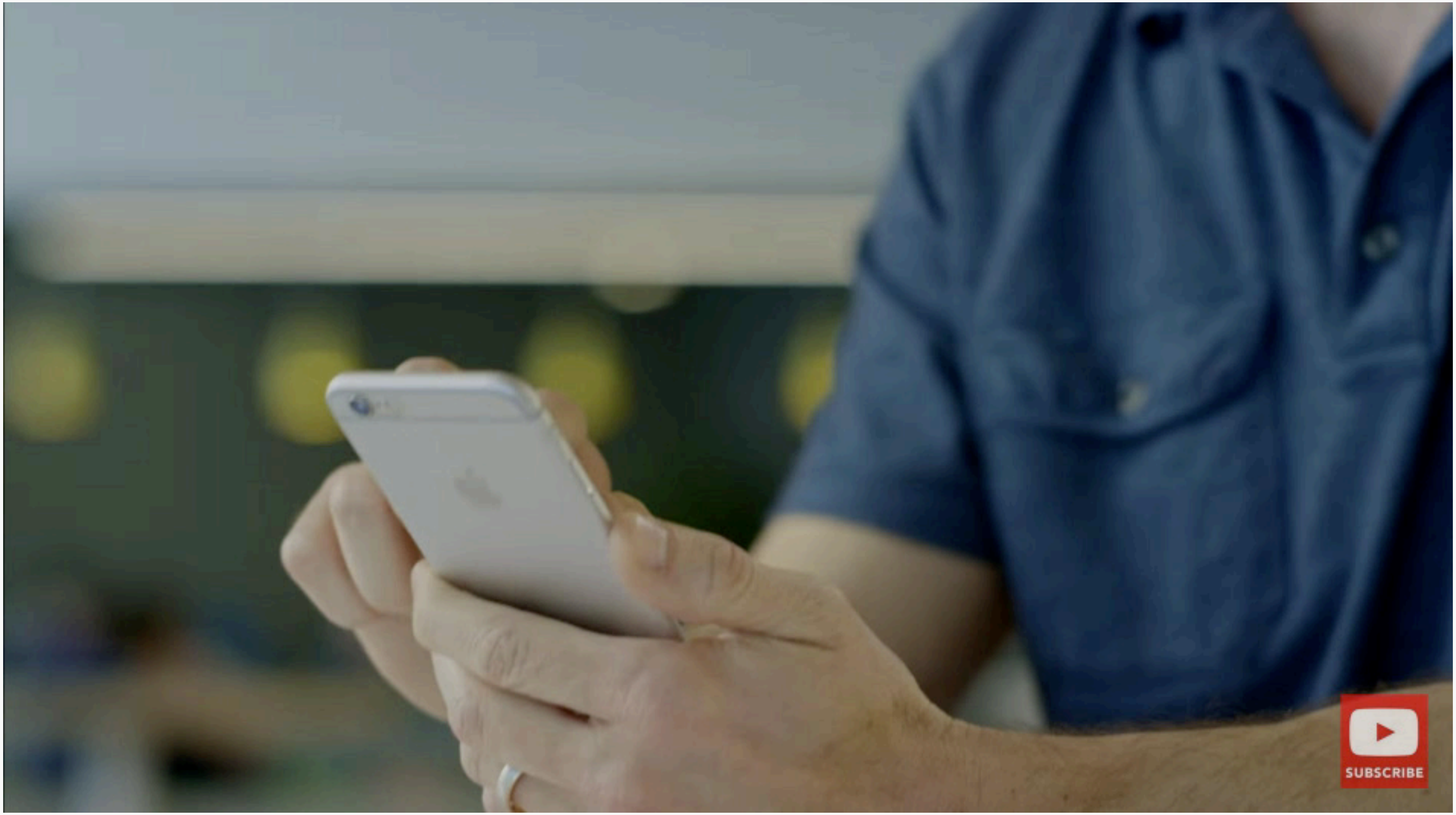
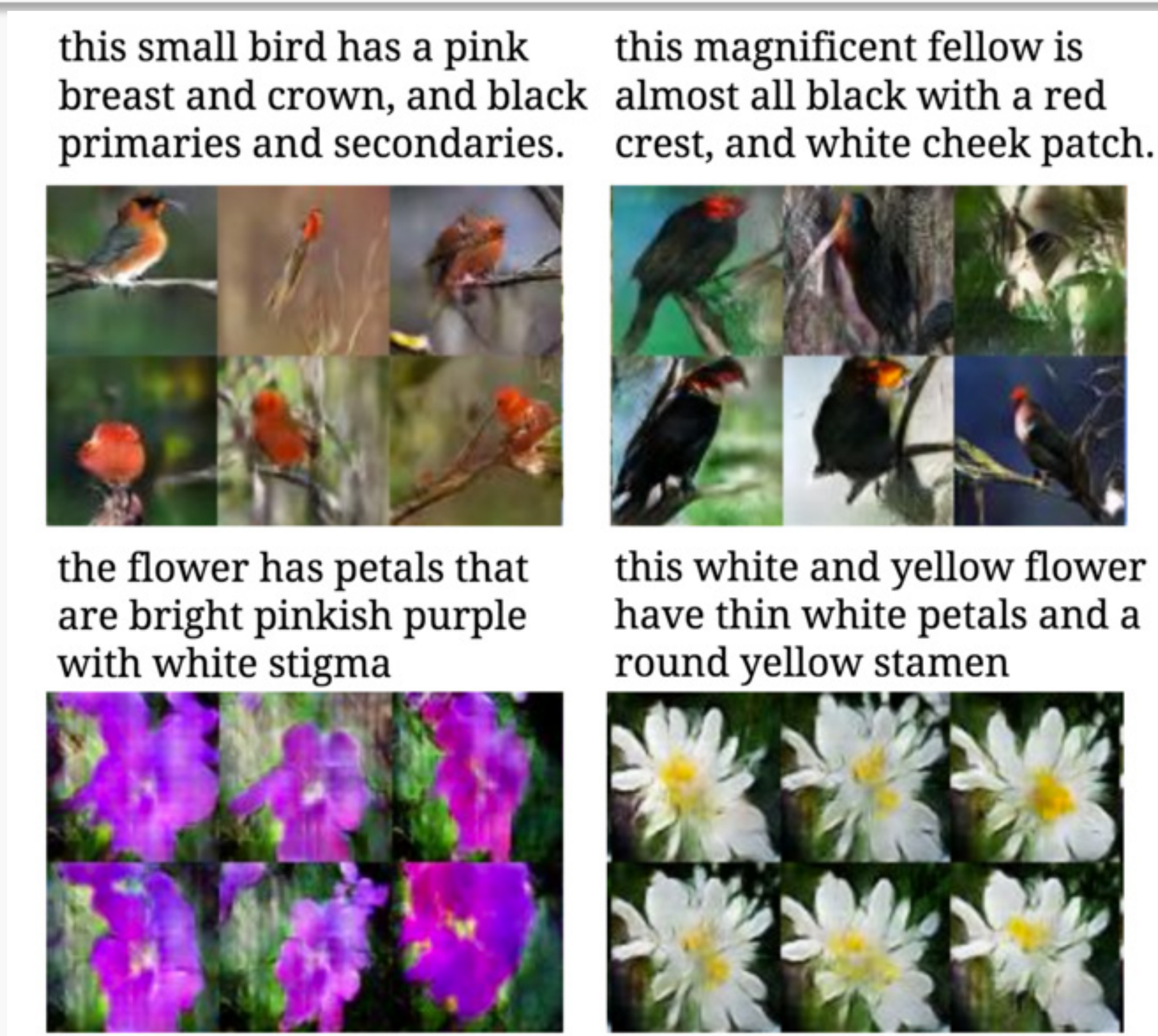


Image Synthesis I

redshank volcano monastery

Anh Nguyen, Jason Yosinski, Yoshua Bengio, Alexey Dosovitskiy, Jeff Clune

Text to Image



Scott Reed, Zeynep Akata, Xinchen Yan, Lajanugen Logeswaran, Bernt Schiele, Honglak Lee

Andrej Karpathy
@karpathy

Folgen

One day we'll be talking about good old "hand-crafted" films and instead the norm will be watching AI-generated (infinite) content on demand

Thank you!

Prof. Dr. Kai Uwe Barthel
barthel@pixolution.de

Hochschule für Technik
und Wirtschaft Berlin

University of Applied Sciences

visual-computing.com

pixolution

pixolution.org