Australian Government

Bureau of Meteorology

A Data Content Search Method

Application of a Pre-Trained Convolution Neural Network for Data Content Similarity Searching within Large Environmental Datasets

Justin Freeman

what?

‘*search this database for data that looks like this*’

‘*find similar events to this event*’

‘*find all days in the dataset that are similar to this day*’

image based content search

methods applied to the task of **finding similar images** that match the content of the **query image**.

image **content** could relate to a specific **object** or the entire **scene**.

search

results

(some) image based content search methods

local sensitivity hashing (**LSH**)

linear discriminant analysis (**LDA**)/principal component analysis (**PCA**)

scale invariant feature transform (**SIFT**)

speed up robust feature (**SURF**)

binary robust independent elementary features (**BRIEF**)

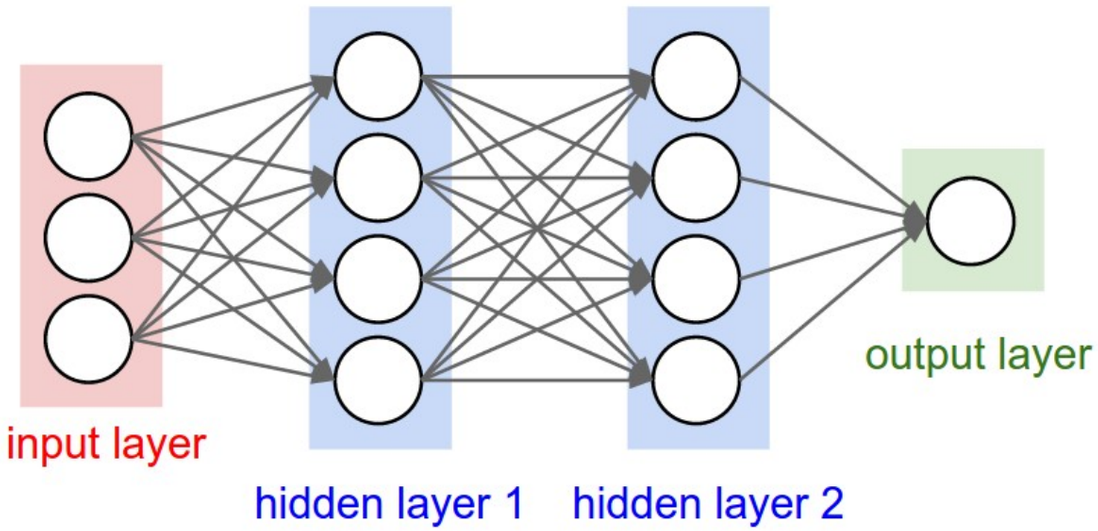
oriented FAST and rotated BRIEF (**ORB**)

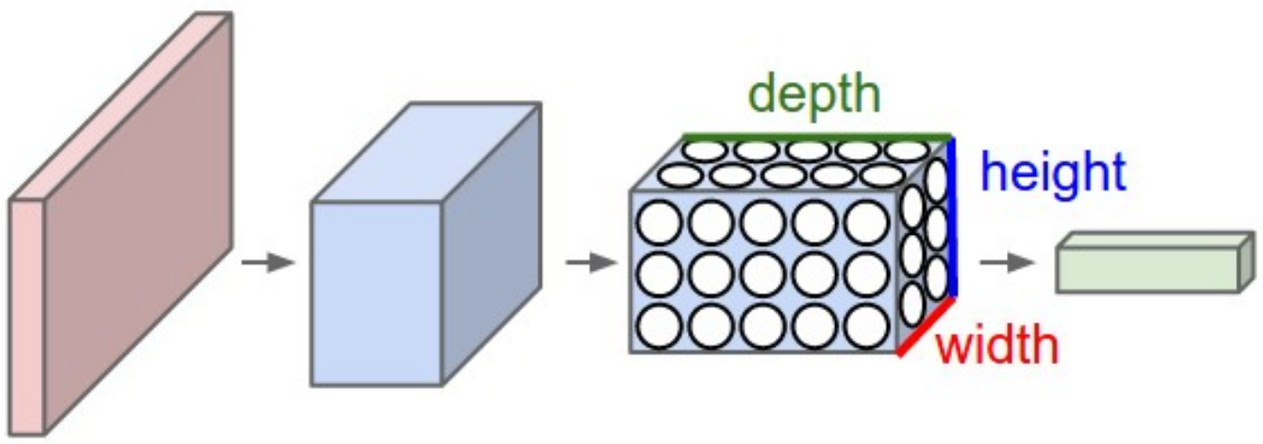
…

convolution neural networks (**CNN**)

…

convolution neural network

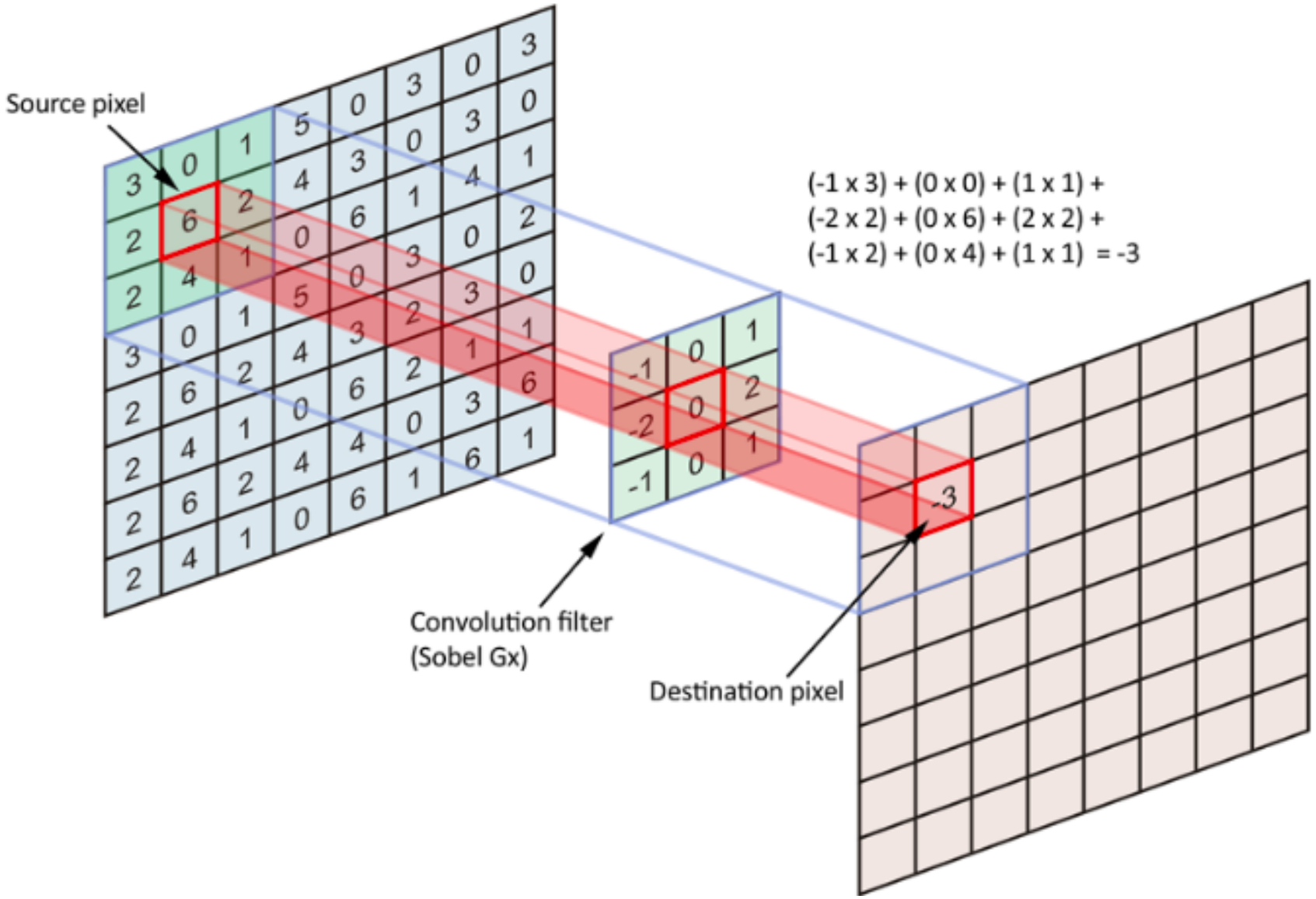




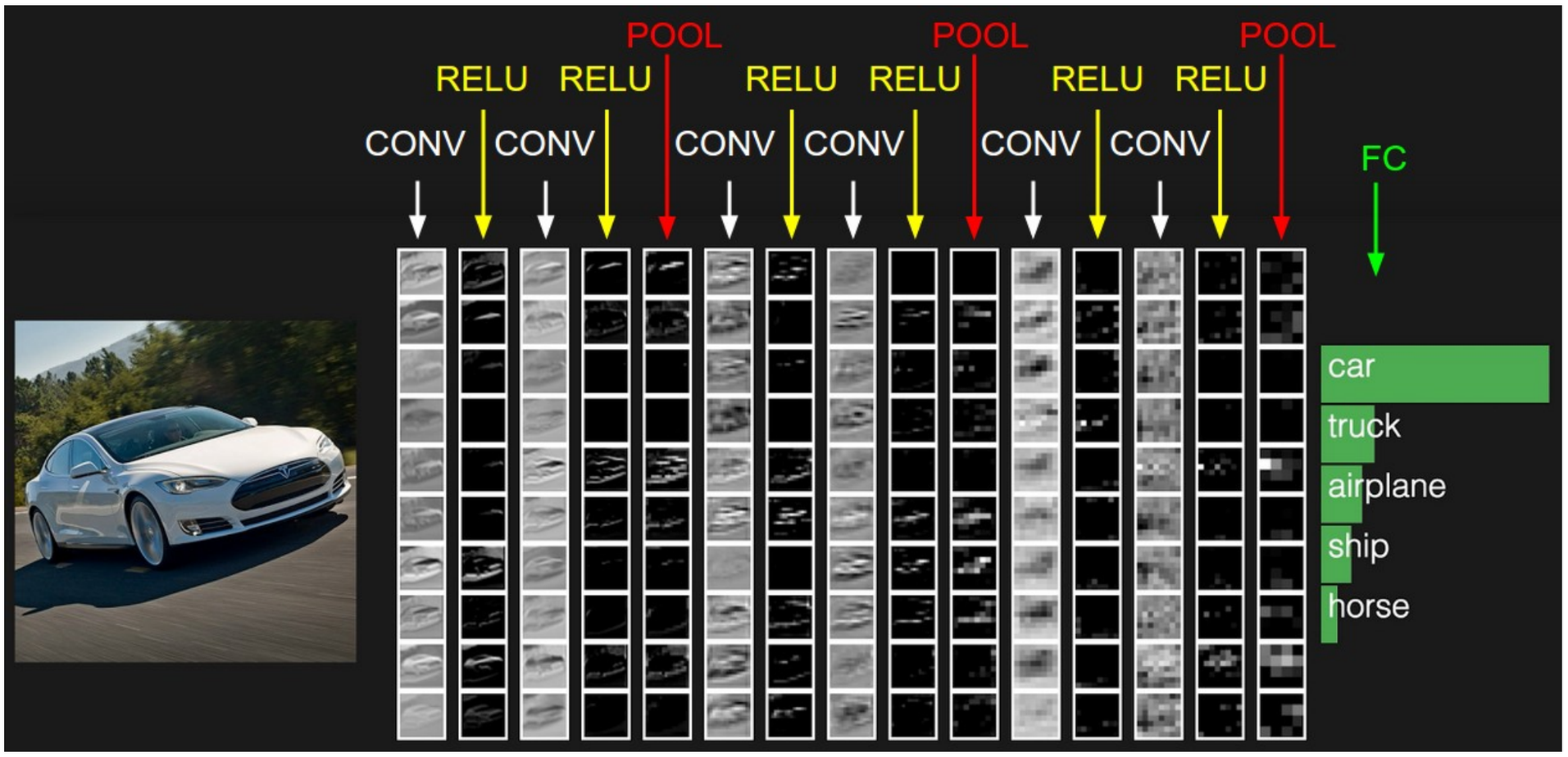
convolution neural networks arrange neurons in 3 dimensions **- width, height, depth**

the neurons in a layer will only be connected to a small region of the layer before it

convolution neural network - filter functions



convolution neural network



convolution neural network - image classification

imagenet (www.image-net.org)

Large Scale Visual Recognition Challenge (ILSVRC*)

This challenge evaluates algorithms for object localization/detection from images/videos at scale.

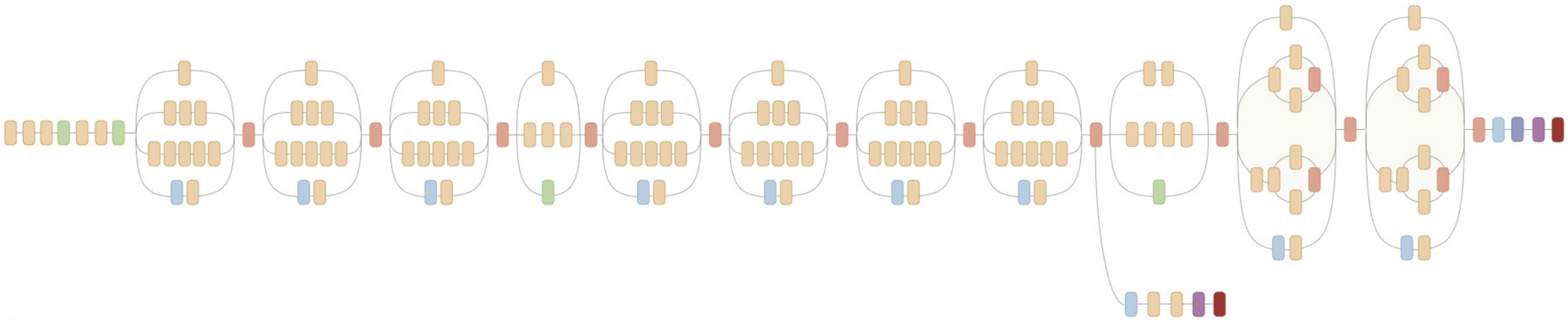
I. Object localization for 1000 categories.
II. Object detection for 200 fully labeled categories.
III. Object detection from video for 30 fully labeled categories.

convolution neural network - image classification

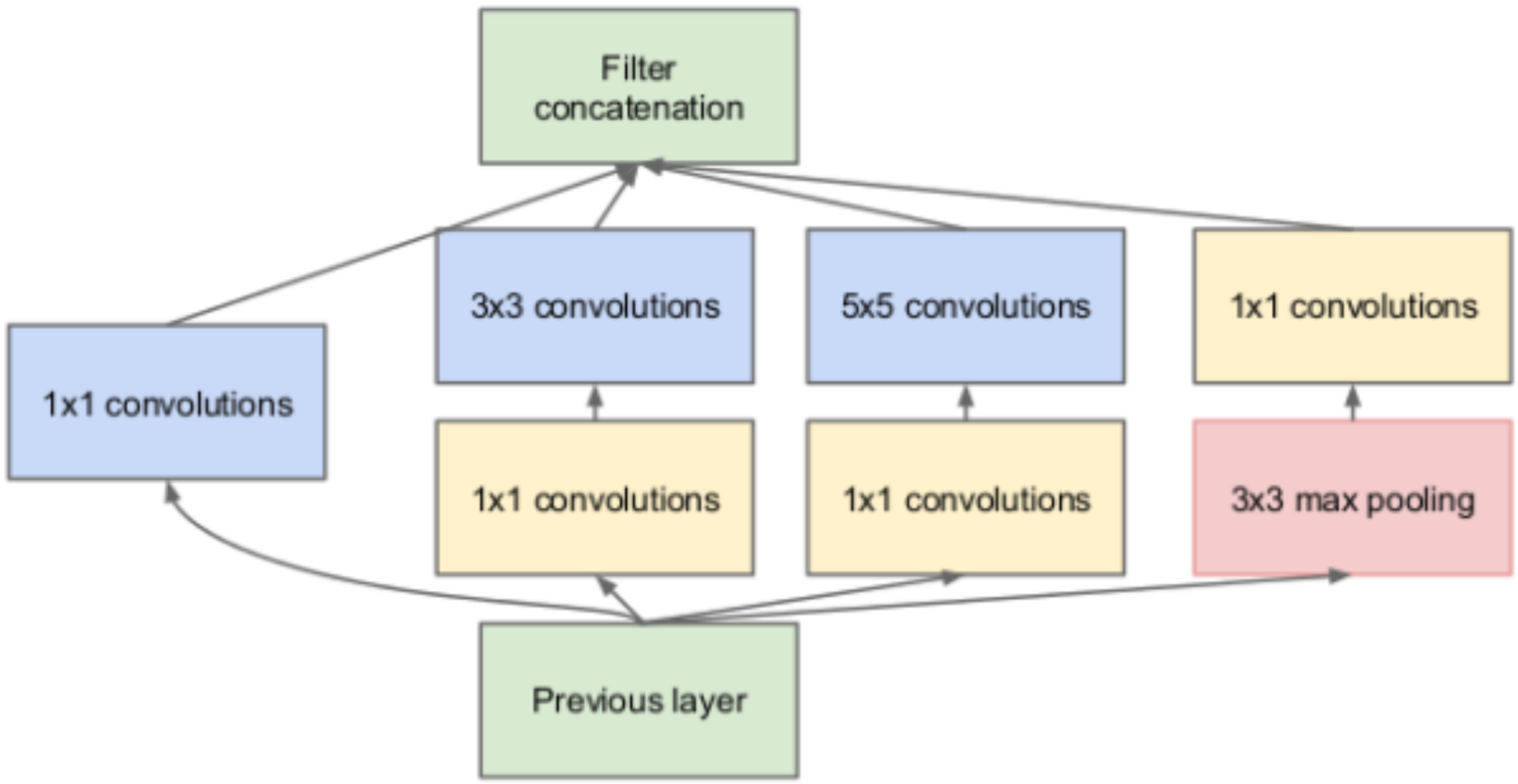
Model	Size	Top-1 Accuracy	Top-5 Accuracy	Parameters	Depth
Xception	88 MB	0.790	0.945	22,910,480	126
VGG16	528 MB	0.713	0.901	138,357,544	23
VGG19	549 MB	0.713	0.900	143,667,240	26
ResNet50	99 MB	0.749	0.921	25,636,712	168
InceptionV3	92 MB	0.779	0.937	23,851,784	159
InceptionResNetV2	215 MB	0.803	0.953	55,873,736	572
MobileNet	16 MB	0.704	0.895	4,253,864	88
MobileNetV2	14 MB	0.713	0.901	3,538,984	88
DenseNet121	33 MB	0.750	0.923	8,062,504	121
DenseNet169	57 MB	0.762	0.932	14,307,880	169
DenseNet201	80 MB	0.773	0.936	20,242,984	201
NASNetMobile	23 MB	0.744	0.919	5,326,716	-
NASNetLarge	343 MB	0.825	0.960	88,949,818	-

inception v3 architecture

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. and Wojna, Z., 2016. **Rethinking the inception architecture for computer vision.** In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2818-2826).



inception module



convolution neural network

can we applying a pre-trained deep convolution neural network to the task of content searching of large environmental data sets?

convolution neural network + content search method

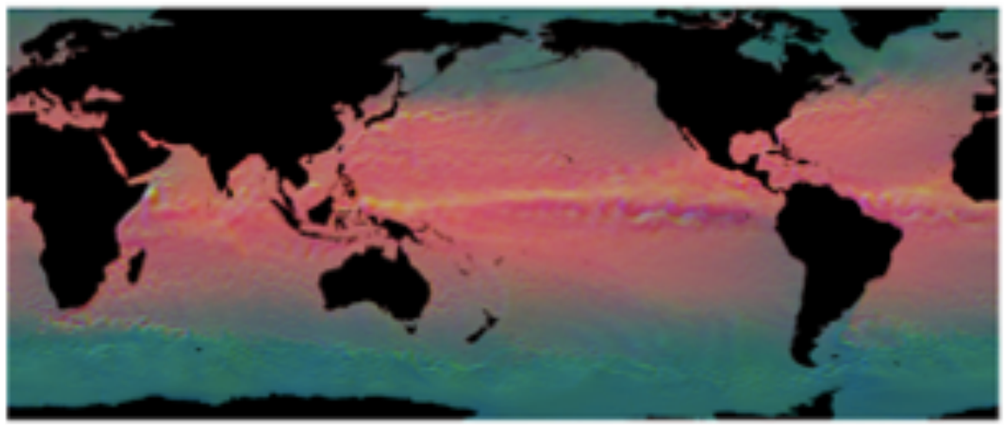
- **extract feature vectors** from the final convolution layer of a pertained inception v3 network.
    - the final fully connected dense layer is omitted from the network architecture.
    - we extract the output from the layer preceding the fully dense layer and store this as a vector representation, or feature map, of the input data.
    - the deeper layers of a CNN represent finer structural components of the input.

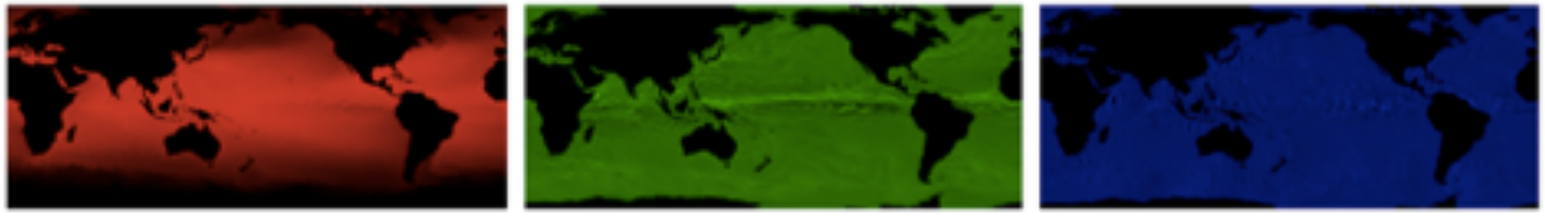
- for the task of content searching**, extract the globally averaged feature maps calculated in the final pooling layer prior to the softmax operation.**
    - The final Inception module outputs 2048 feature maps of size 8x8.
    - With a patch size of 8x8, the average pooling layer reduced the dimension from 8x8x2048 to 1x1x2048 by globally averaging each feature map.
    - For each input field, we extract and store the vector of coefficients from the average pooling operation, yielding a collection of vectors which describe each instance.

data preprocessing

- for each reanalysis field we construct a **stacked image representation** of the data.
  - for ocean data (BRAN, Oke et al., 2012), we map the
    - **sea surface temperature** to the red channel,
    - and the **u** and **v** components of surface velocity to the green and blue channel.

  - for the atmospheric data (JRA-55, Ebita et al., 2011),
    - the red channel contains the **10 metre meridional wind**
    - the green channel contains the **10 metre zonal wind** components.
    - the blue channel of the generated image file is set to a pixel value of **zero**.





sst **u** **v**

similarity metric

Content search of new data field to find a similar field within the precomputed feature map dataset is undertaken using a ***k* nearest neighbour search** method.

***k* nearest neighbours** is a proximity search method which will find the *k* closest, or most similar, data points in a data set to a given point.

The distance metric is defined as

$$\cos\theta = \frac{\overrightarrow{a} \cdot \overrightarrow{b}}{\| \overrightarrow{a} \| \| \overrightarrow{b} \|}$$

and the similarity metric is defined as

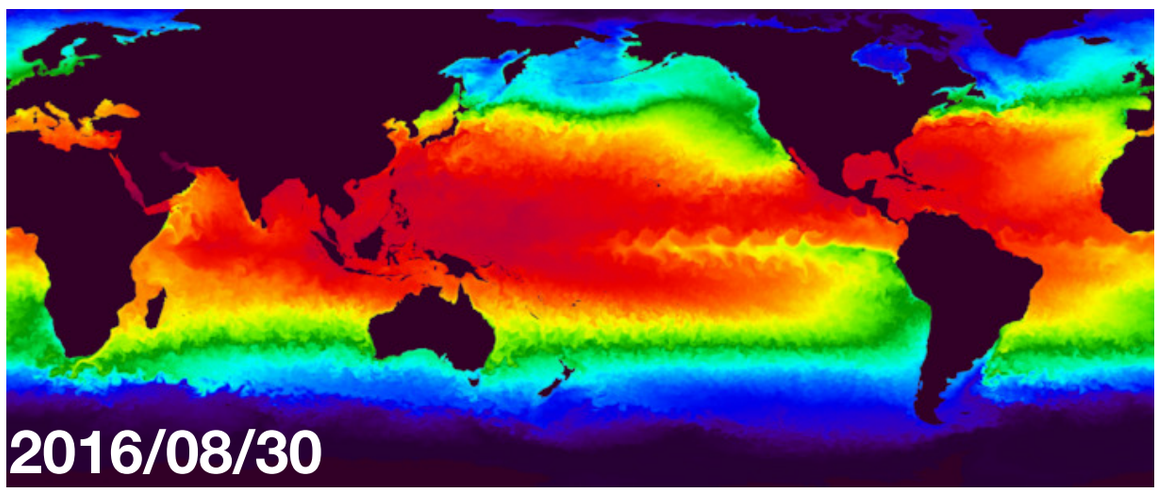
$$\textbf{similarity} = 1 - \textbf{distance}$$

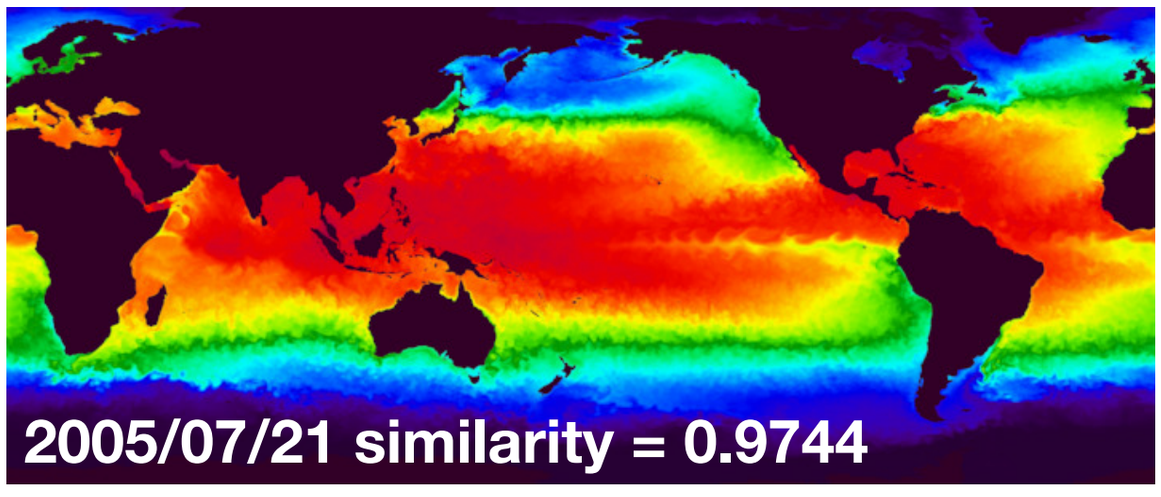
where

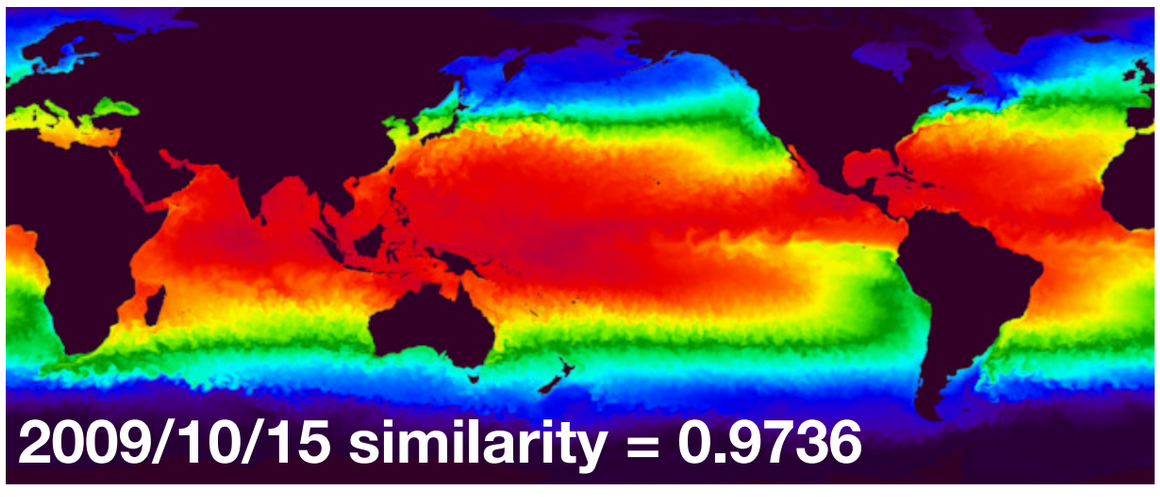
- **0** indicates that the fields **have nothing in common**
- **1** indicates **very similar fields**

results - BRAN content search

search

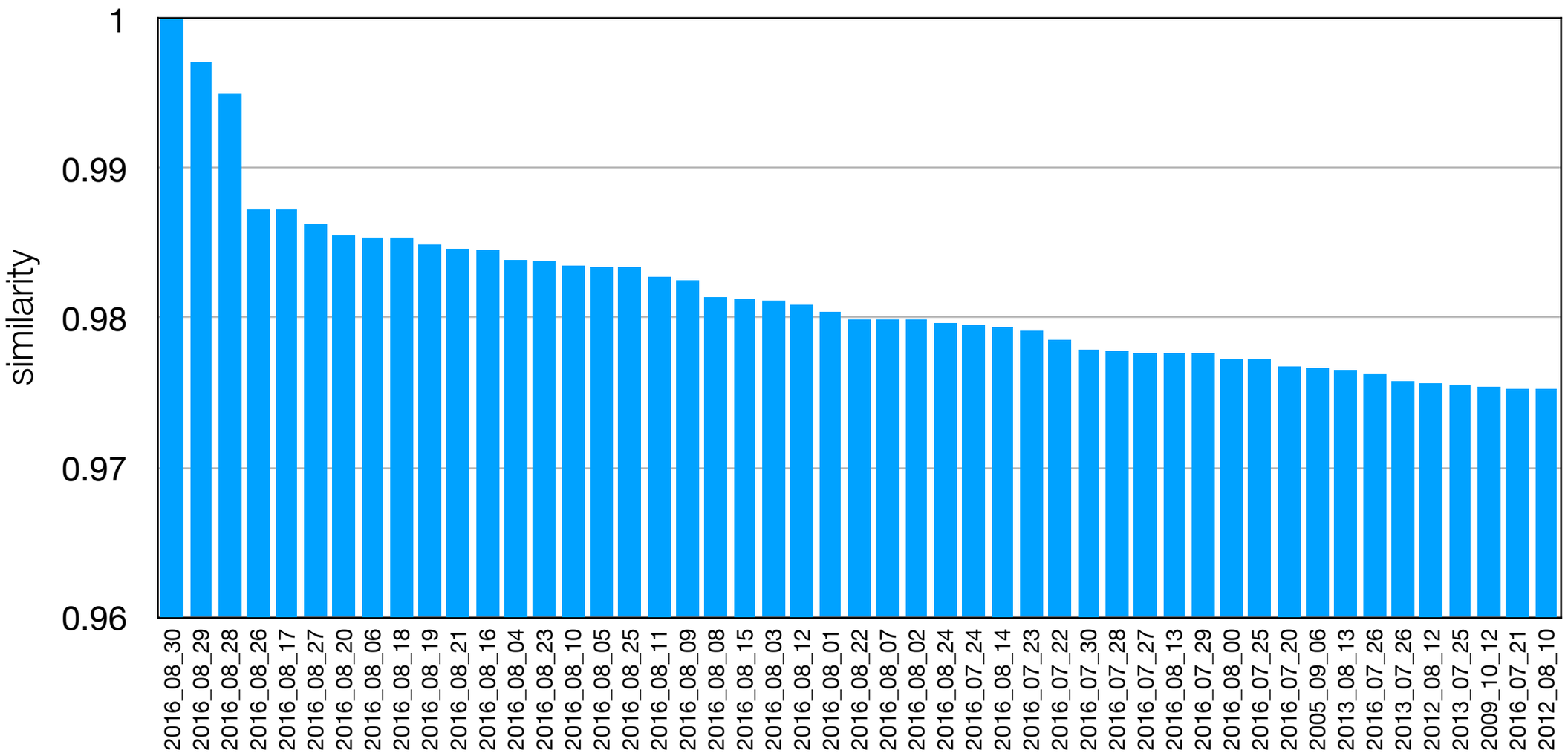




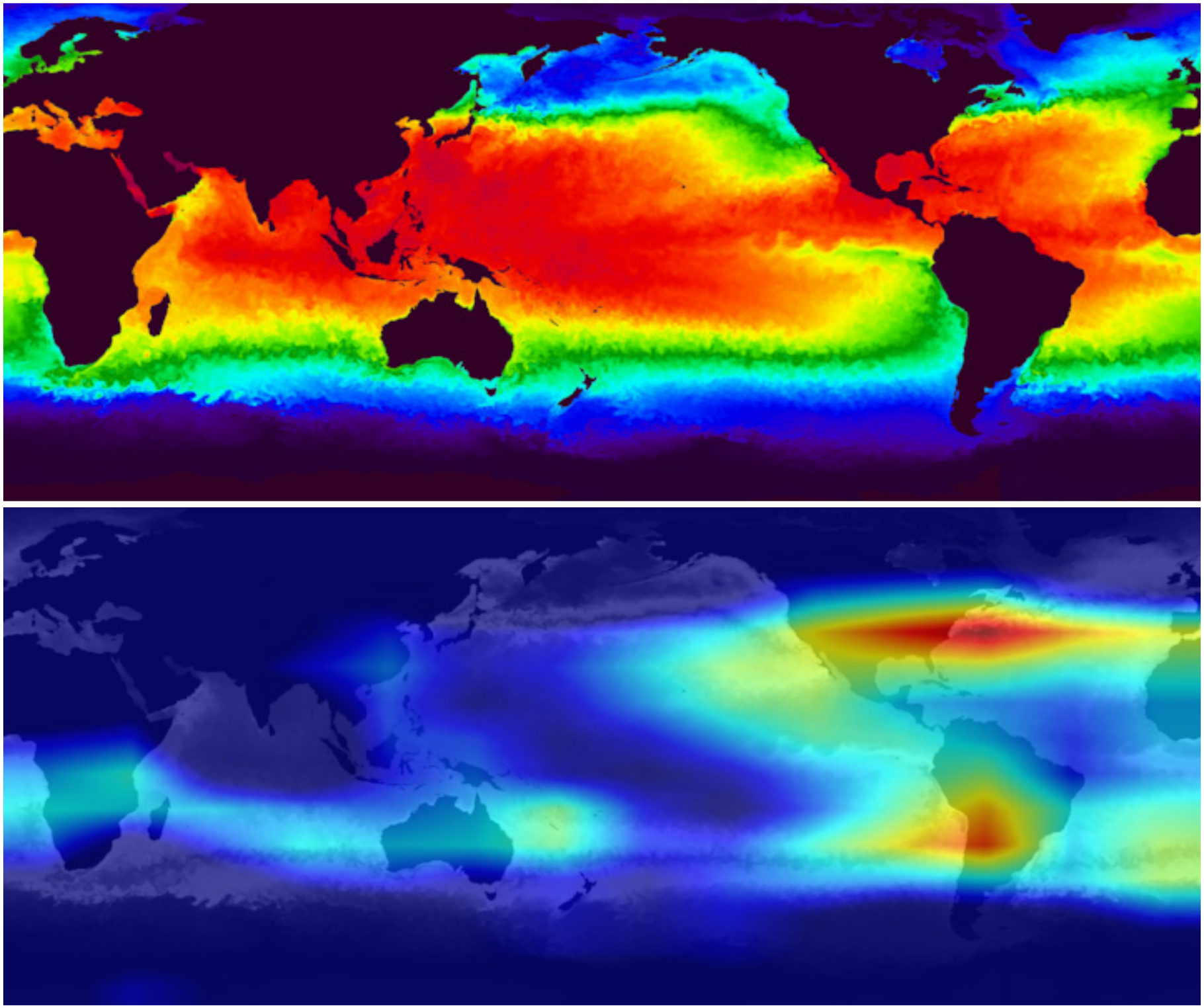


similar fields

results - BRAN content search

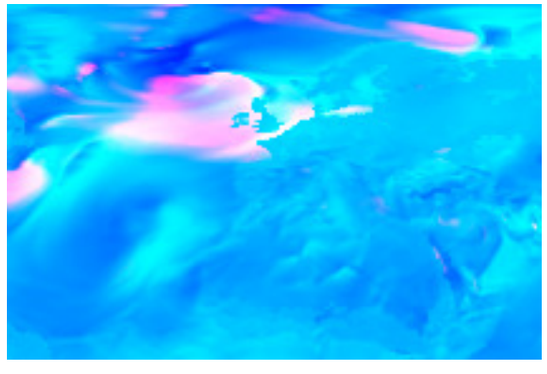


CNN attention map

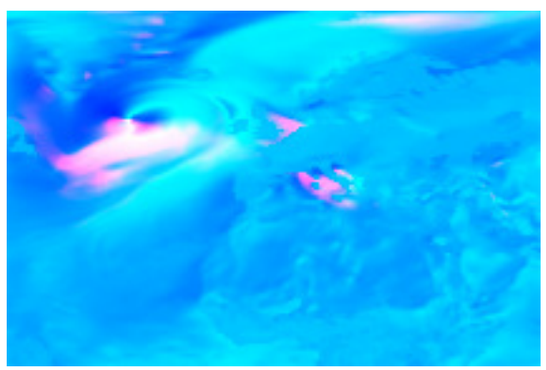


results - JRA55 content search (Ebita et al., 2011)

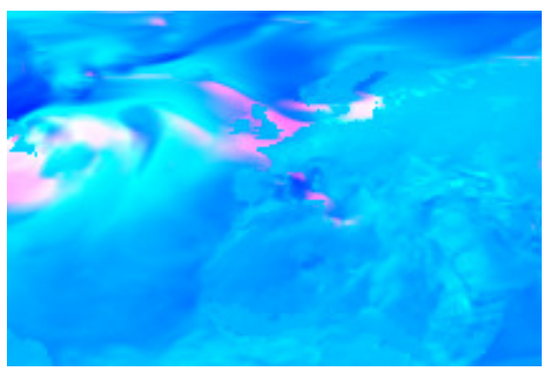
28/1/200



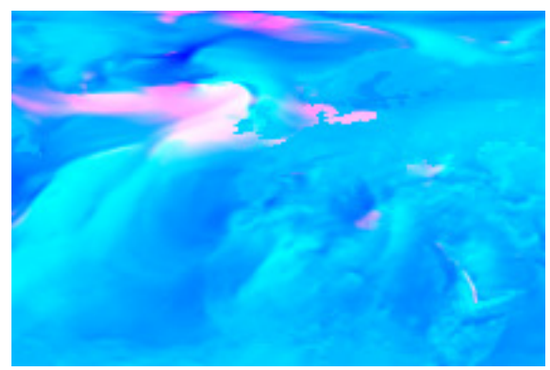
2/10/2017, similarity = 0.9388

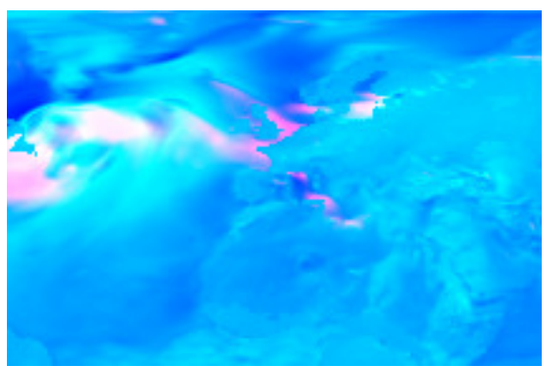


29/12/2011, similarity = 0.9368



21/2/2017, similarity = 0.9353





30/12/2011, similarity = 0.9362

applications

content search + find similar
ensemble generation
ensemble dressing
anomaly search
extreme events
forecast guidance
ensemble modelling + initialisation
…
natural language search
automatic forecast text generation
…