High Performance Image Search & Retrieval in Remote Sensing Data

Sheik Mohamed Imran
AI Technical Solutions Specialist, Intel
14 March 2022

Notices & Disclaimers

INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL® PRODUCTS. EXCEPT AS PROVIDED IN INTEL'S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, INTEL ASSUMES NO LIABILITY WHATSOEVER, AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY RELATING TO SALE AND/OR USE OF INTEL PRODUCTS, INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT, OR OTHER INTELLECTUAL PROPERTY RIGHT. Intel products are not intended for use in medical, life-saving, life-sustaining, critical control or safety systems, or in nuclear facility applications.

Intel products may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.

This document may contain information on products in the design phase of development. The information herein is subject to change without notice. Do not finalize a design with this information. Intel may make changes to dates, specifications, product descriptions, and plans referenced in this document at any time, without notice.

Designers must not rely on the absence or characteristics of any features or instructions marked "reserved" or "undefined." Intel reserves these for future definition and shall have no responsibility whatsoever for conflicts or incompatibilities arising from future changes to them.

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Learn more at intel.com, or from the OEM or retailer.

No computer system can be absolutely secure.

Intel Corporation or its subsidiaries in the United States and other countries may have patents or pending patent applications, trademarks, copyrights, or other intellectual property rights that relate to the presented subject matter. The furnishing of documents and other materials and information does not provide any license, express or implied, by estoppel or otherwise, to any such patents, trademarks, copyrights, or other intellectual property rights.

Performance estimates or simulated results based on internal Intel analysis or architecture simulation or modeling are provided to you for informational purposes. Any differences in your system hardware, software or configuration may affect your actual performance.

Performance tests and ratings are measured using specific computer systems and/or components and reflect the approximate performance of Intel products as measured by those tests. Any difference in system hardware or software design or configuration may affect actual performance. Buyers should consult other sources of information to evaluate the performance of systems or components they are considering purchasing. For more complete information about performance and benchmark results, visit www.intel.com/benchmarks.

Intel does not control or audit third-party benchmark data or the web sites referenced in this document. You should visit the referenced web site and confirm whether referenced data are accurate.

Optimization Notice: Intel's compilers may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel microprocessors. These optimizations include SSE2, SSE3, and SSSE3 instruction sets and other optimizations. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel. Microprocessor-dependent optimizations in this product are intended for use with Intel microprocessors. Certain optimizations not specific to Intel microarchitecture are reserved for Intel microprocessors. Please refer to the applicable product User and Reference Guides for more information regarding the specific instruction sets covered by this notice. Notice Revision #20110804.

Intel® Advanced Vector Extensions (Intel® AVX)* provides higher throughput to certain processor operations. Due to varying processor power characteristics, utilizing AVX instructions may cause a) some parts to operate at less than the rated frequency and b) some parts with Intel® Turbo Boost Technology 2.0 to not achieve any or maximum turbo frequencies. Performance varies depending on hardware, software, and system configuration and you can learn more at http://www.intel.com/go/turbo.

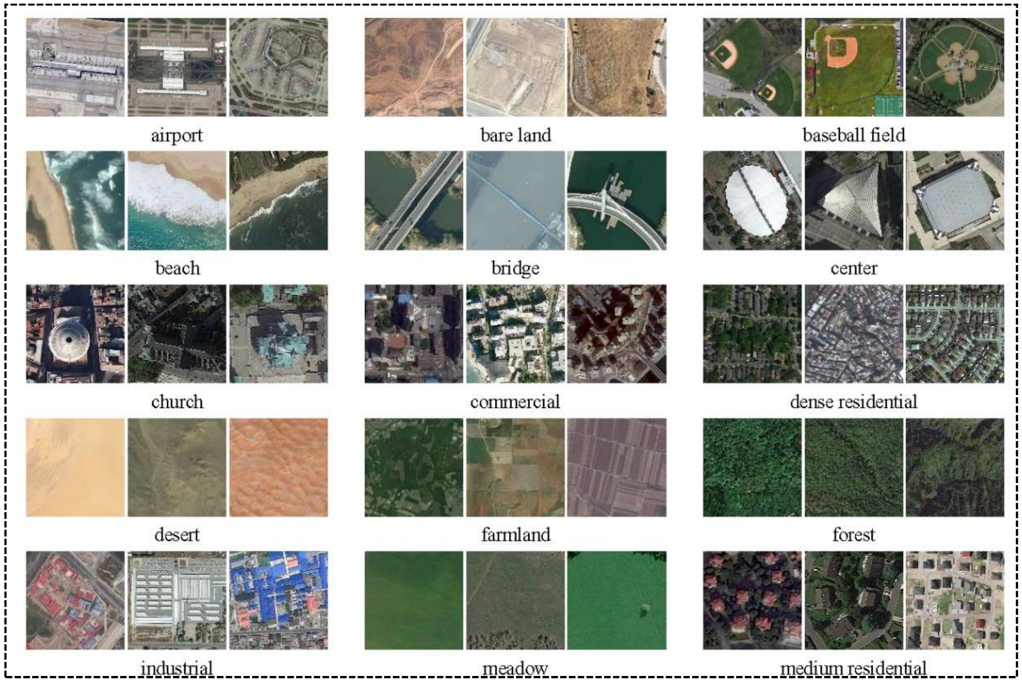
Code names are used by Intel to identify products, technologies, or services that are in development and not publicly available. These are not "commercial" names and not intended to function as trademarks.

Intel processors of the same SKU may vary in frequency or power as a result of natural variability in the production process

What is Remote Sensing Data?

Remote Sensing Image Scene Classification (RESISC) Dataset

Remote Sensing Data



AID (Aerial Image Dataset) - 10000 images within 30 classes, extracted from Google Earth

Source: [link]

Remote Sensing Data

LoveDA: A Remote Sensing Land-Cover Dataset for Domain Adaptive Semantic Segmentation

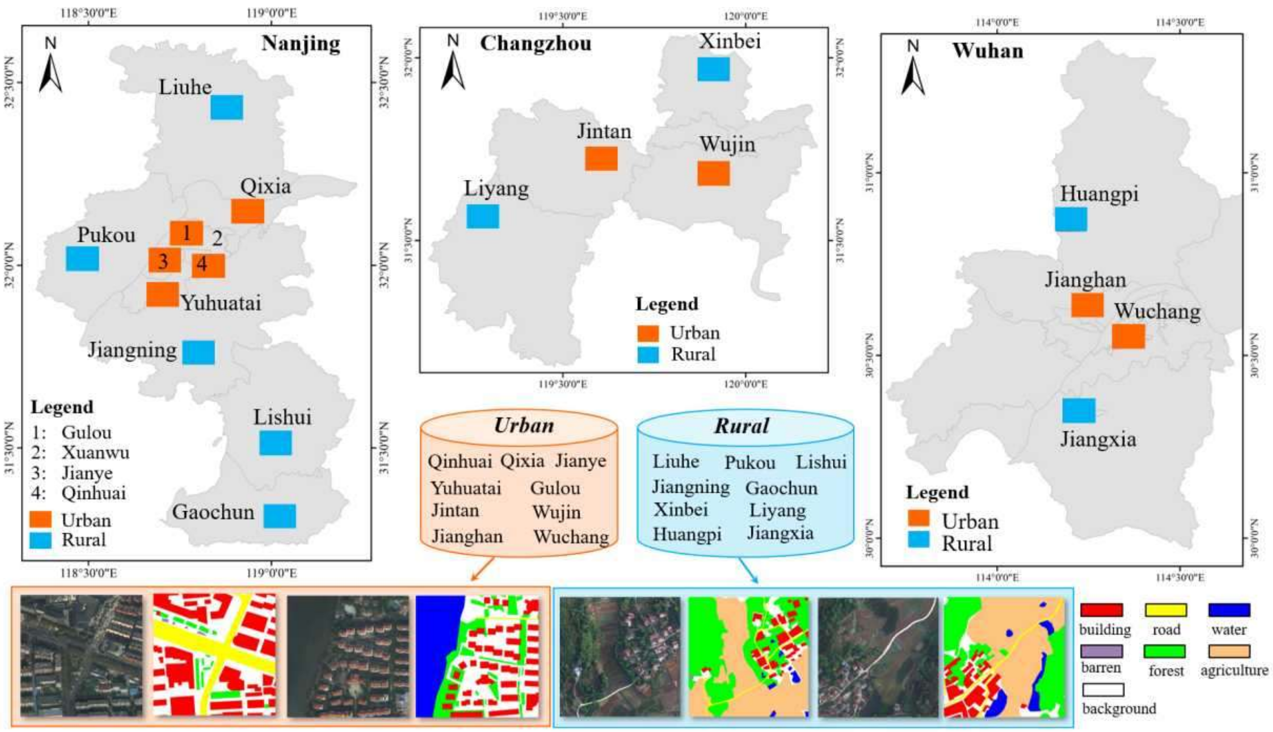
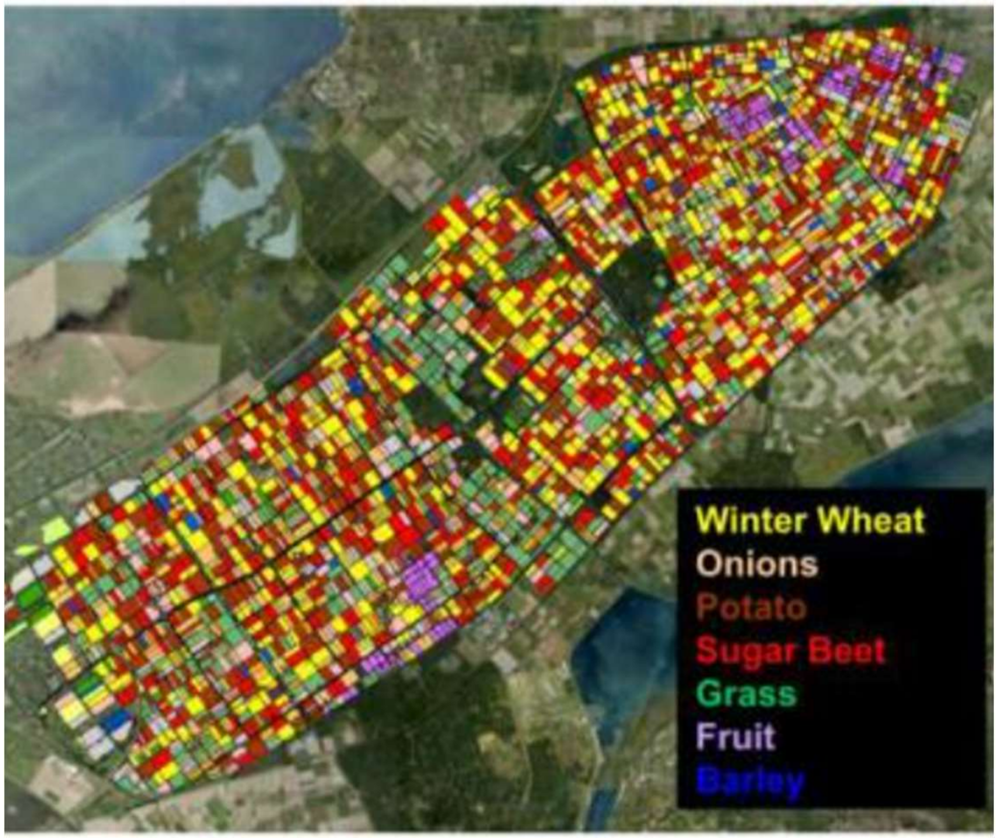


Figure 1: Overview of the dataset distribution. The images were collected from Nanjing, Changzhou, and Wuhan cities, covering 18 different administrative districts.

[Source]

Remote Sensing Data



AgriSAR 2009: C-band SAR data from the Sentinel-1 mission.

Used for agricultural crops dominated/non-dominated

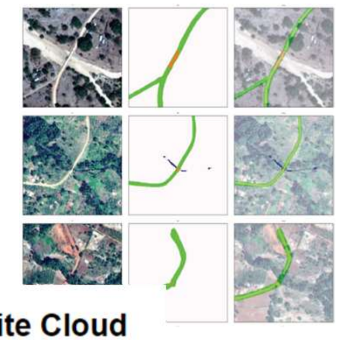
Extensive crop maps as well as crop details for selected representative fields

Source

Remote Sensing Data – Use Cases

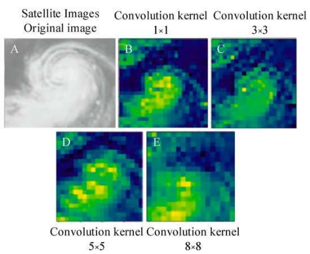
- Detecting weather events
  - Weather nowcasting
- Land Cover Mapping
  - Land resource survey
  - Spatial Planning
  - Land utilization /change detection
  - Understanding urban/rural areas
    - Road extraction, Building extraction
  - Vegetation Detection

Intel & American Red Cross Southern Uganda Bridge Identification



Classification and Prediction of Typhoon Levels by Satellite Cloud Pictures through GC–LSTM Deep Learning Model

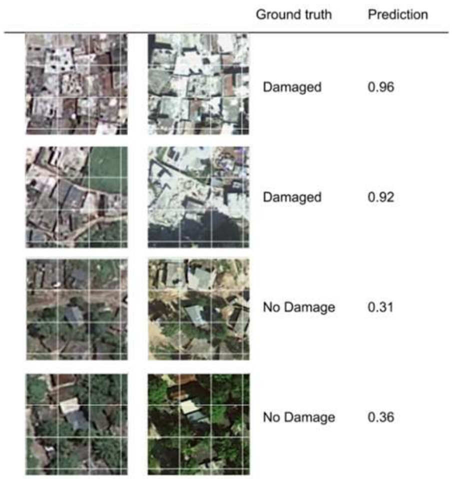
by Jianyin Zhou [1], Jie Xiang [1,*] and Sixun Huang [1,2]



Machine Learning-based Damage Assessment for Disaster Relief

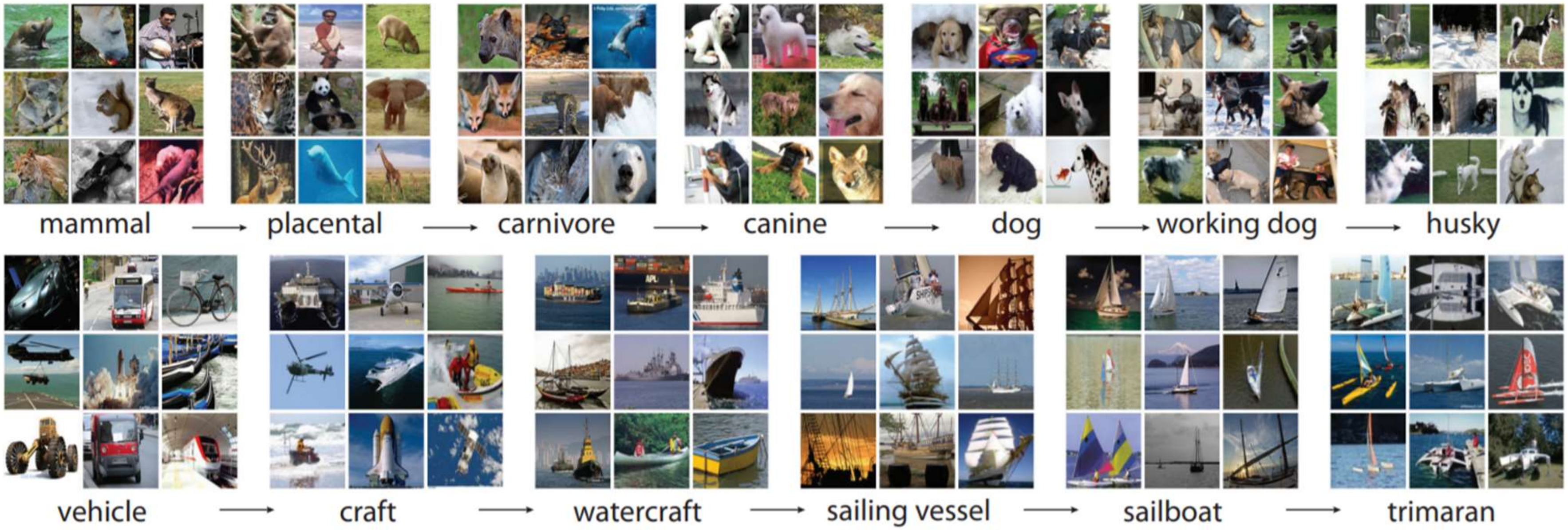
Tuesday, June 16, 2020

Posted by Joseph Xu, Senior Software Engineer and Pranav Khaitan, Engineering Lead, Google Researc



ImageNet

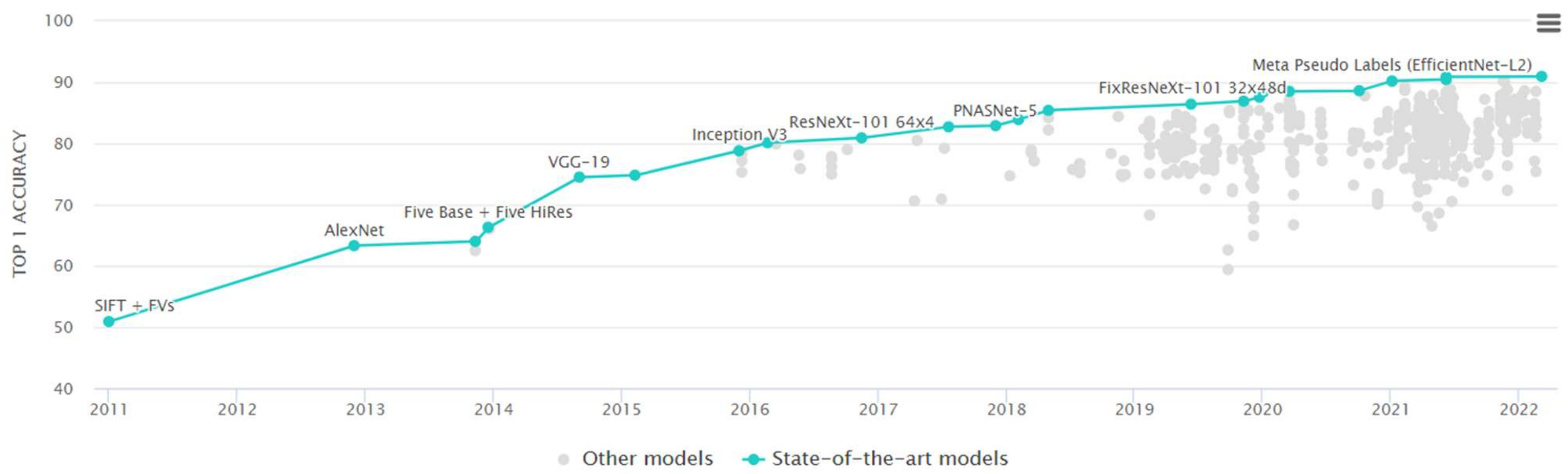
Categories of Images available in ImageNet



[Source](#)

Image Classification on ImageNet

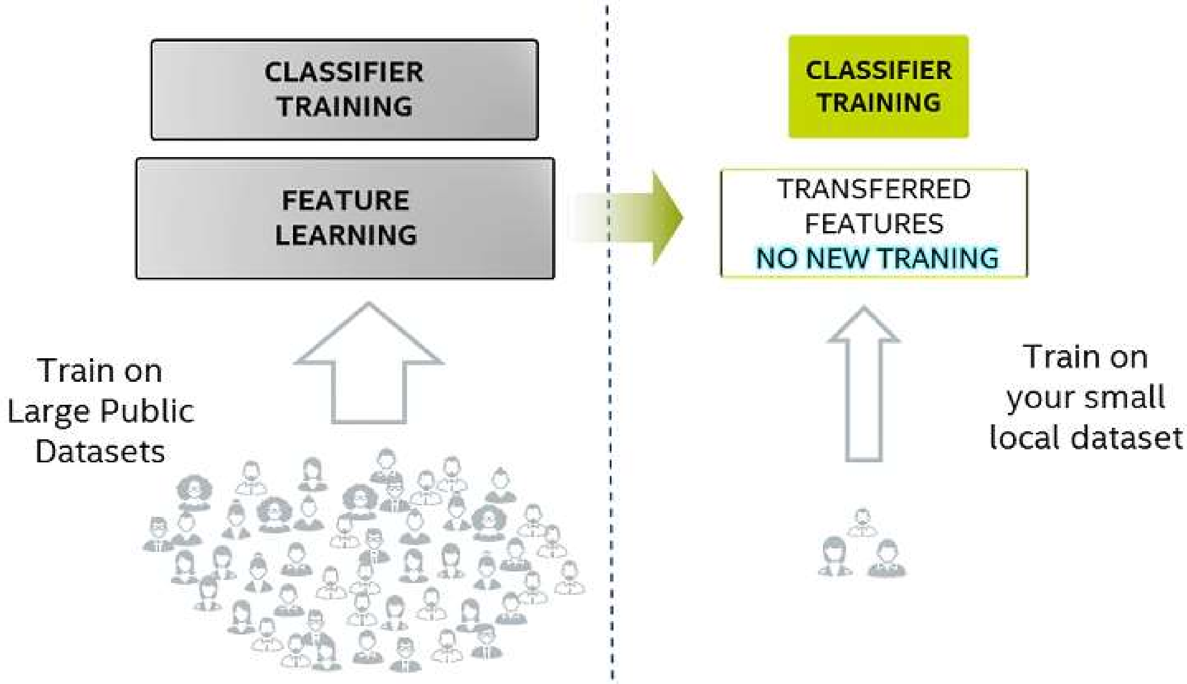
Performance of various models over years



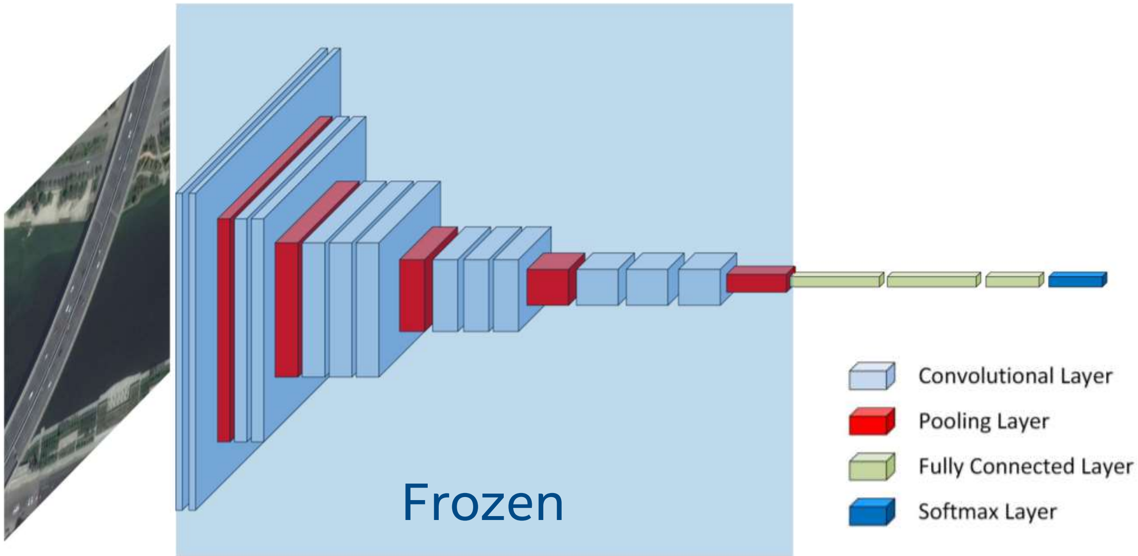
[Source]

Transfer Learning

Re-using the weights from pre-trained models



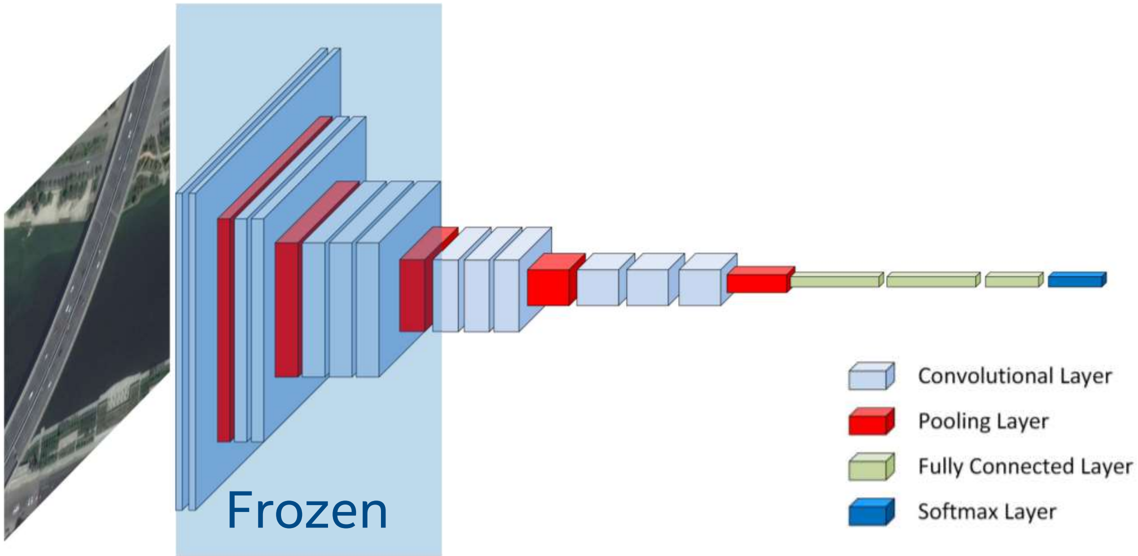
Transfer Learning - Strategies



- Use pre-trained model for Feature extraction
- To be used if custom data similar to data used in base model

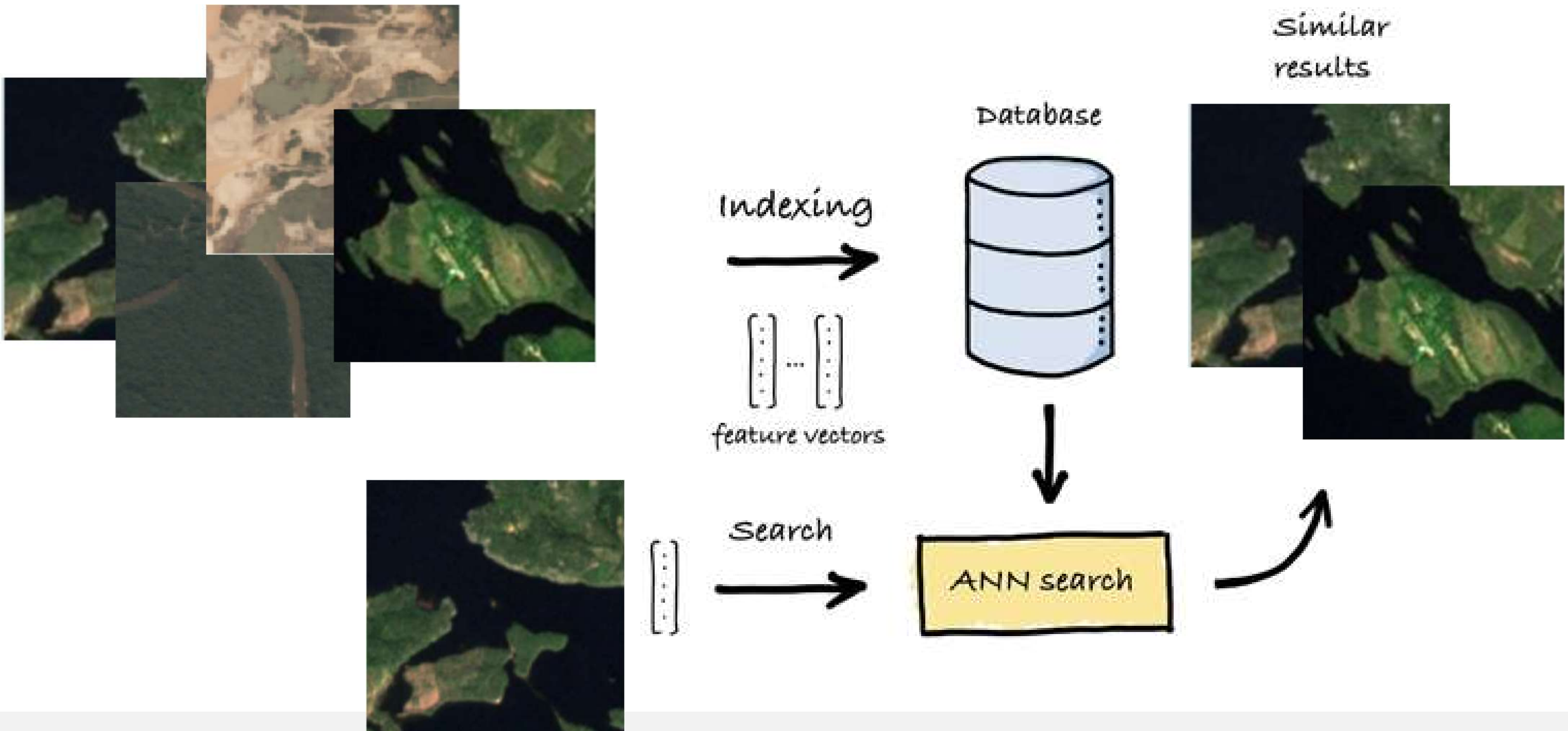
Transfer Learning - Strategies

Transfer Learning Strategies

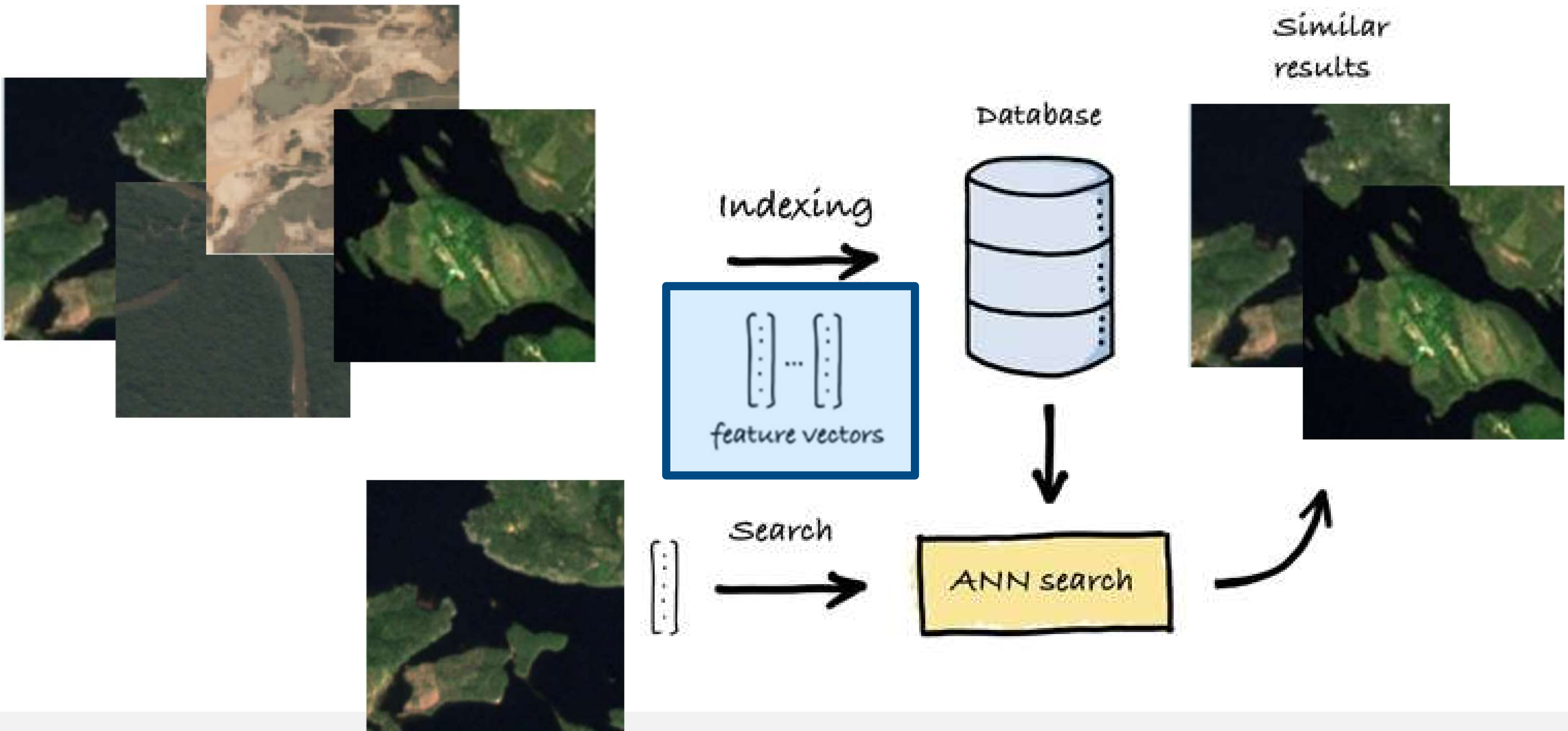


- Use initial layers of pre-trained model for Feature extraction
- To be used if custom data dis-similar to data used in base model

Content Based Image Retrieval System

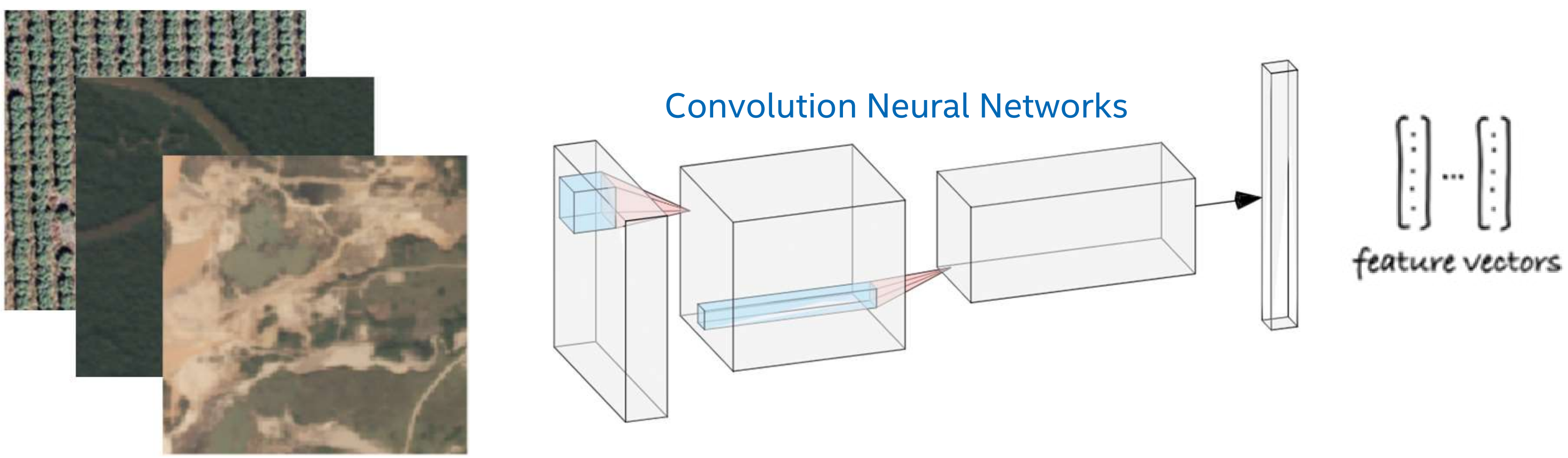


Creating Feature Vectors



Creating Feature Vectors

- Images passed through CNNs to get feature vectors
- These are low dimensionality representation of the input images



Creating Feature Vectors

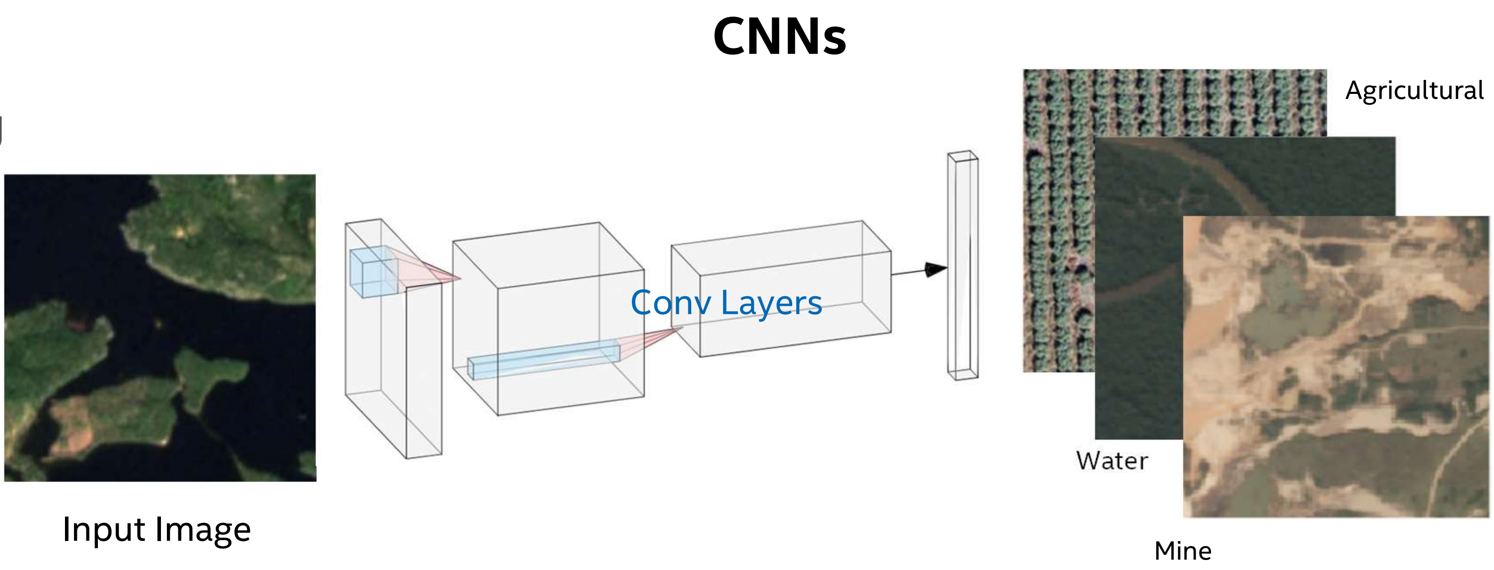
Popular learning techniques to train models for feature extraction

- Classification
- Reconstruction
- Metric Learning

Creating Feature Vectors

Popular learning techniques to train models for feature extraction

- Classification
- Reconstruction
- Metric Learning

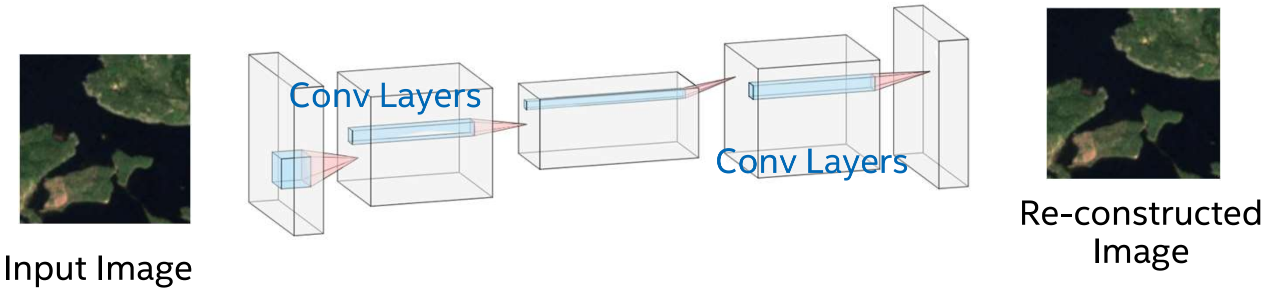


Creating Feature Vectors

Popular learning techniques to train models for feature extraction

- Classification
- Reconstruction
- Metric Learning

Auto Encoders

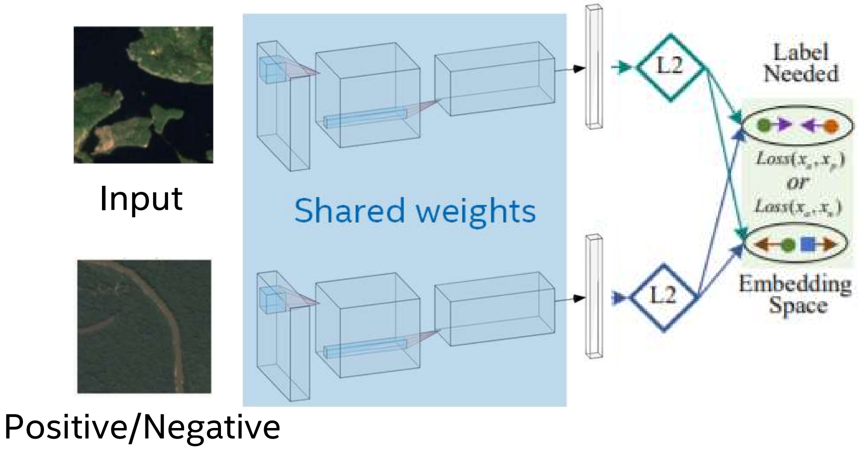


Creating Feature Vectors

Popular learning techniques to train models for feature extraction

- Classification
- Reconstruction
- Metric Learning

Contrastive Loss

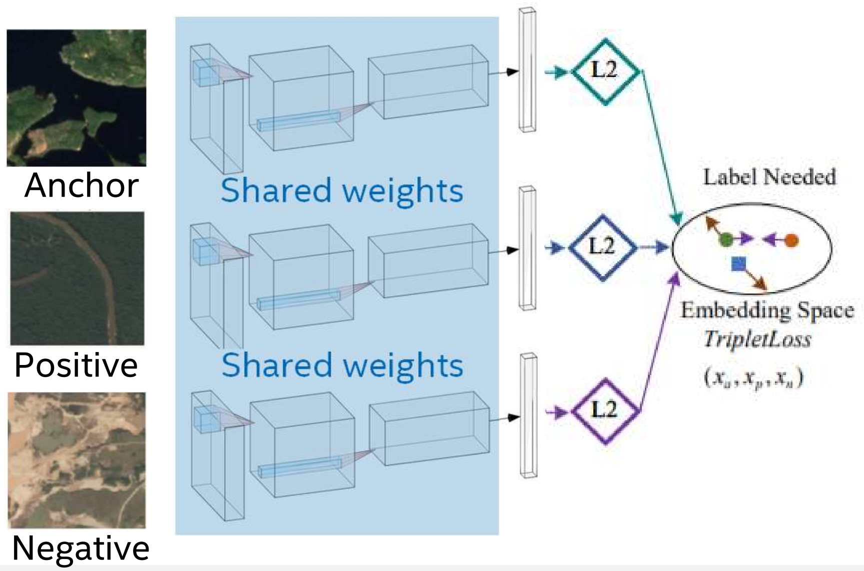


Creating Feature Vectors

Popular learning techniques to train models for feature extraction

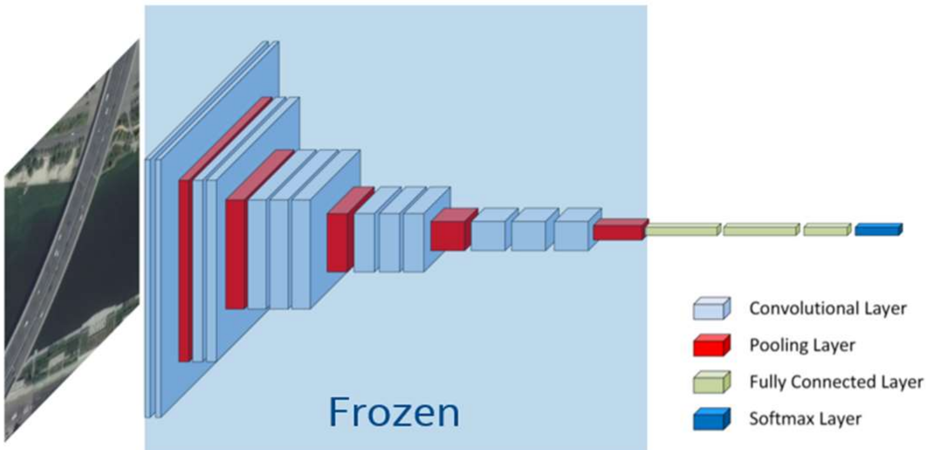
- Classification
- Reconstruction
- Metric Learning

Triplet Loss



Creating Feature Vectors

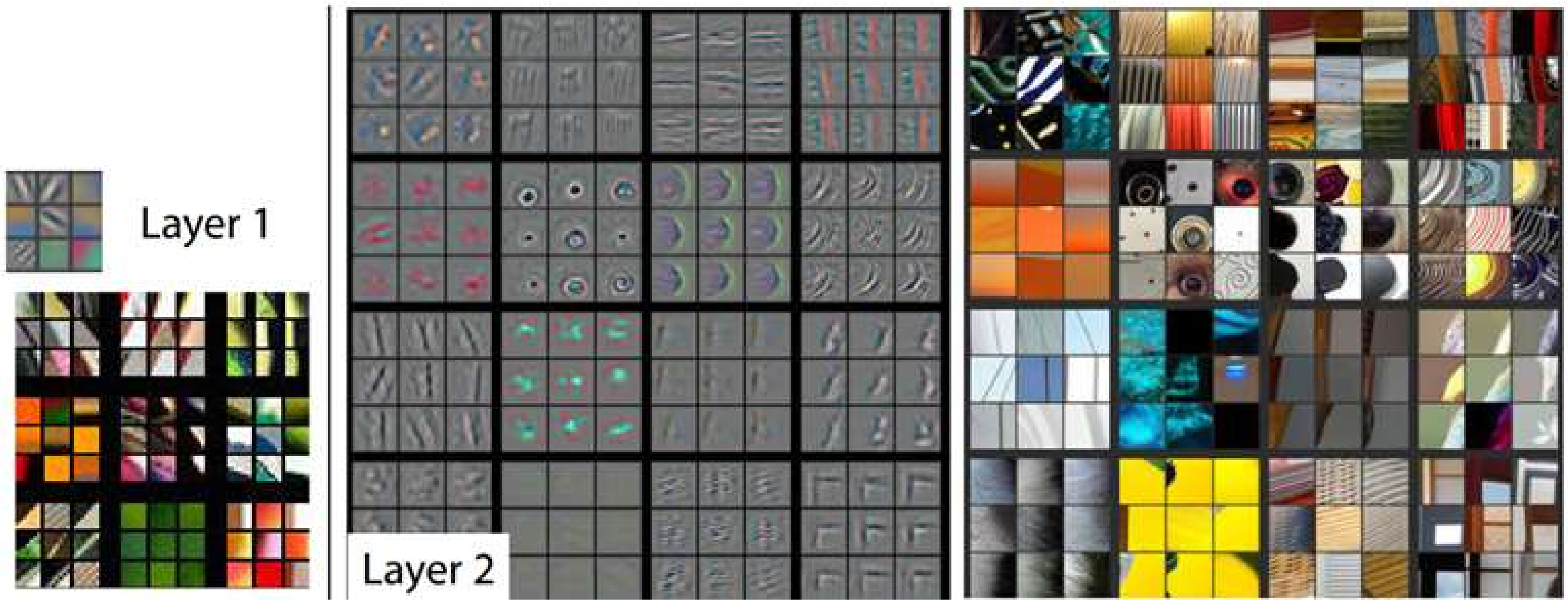
- Supervised Learning (Pre-trained weights ??)



Learning Strategy	Annotated Images
Classification	Yes
Reconstruction	No
Metric Learning (Contrastive)	Yes
Metric Learning (Triplet)	Yes

Creating Feature Vectors

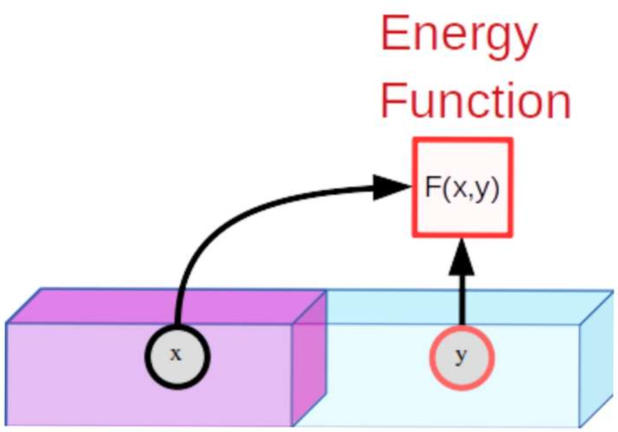
- Supervised Learning (Pre-trained weights ??)



Visualization of features in a fully trained model [4]

Creating Feature Vectors

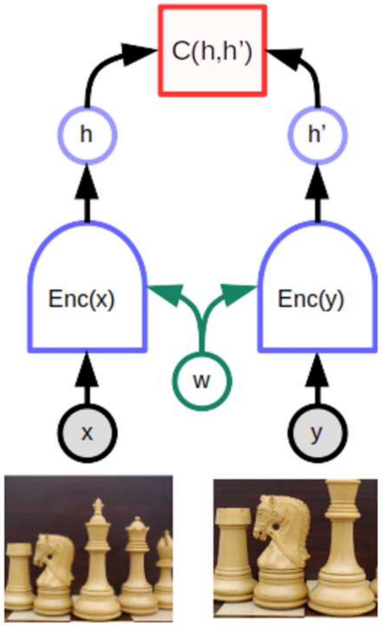
- Energy-based models (EBM)
  - Measures the compatibility between an observation x and a proposed prediction y.
  - If x and y are compatible, the energy is a small number; else it is a larger number.



Source

Creating Feature Vectors

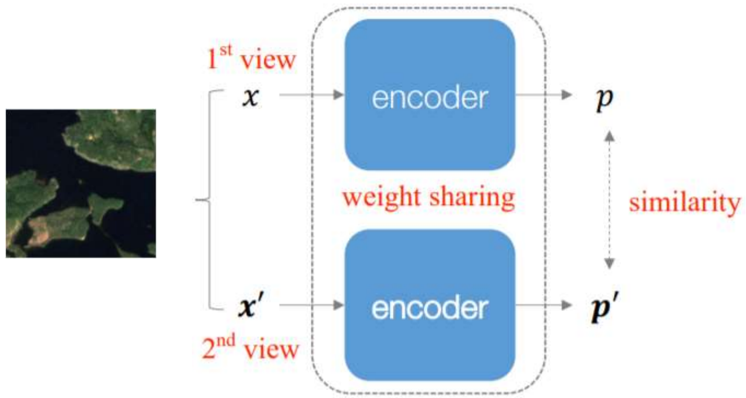
- Energy-based models (EBM)
  - Siamese networks are Energy based models.
  - Uses joint embedding architecture (two copies of same network)in core.
  - Energy is captured as distance between the output vectors.
  - The difficult part is to train the model so that it produces high energy (i.e., different embeddings) for images that are different.



Joint embedding architecture

Creating Feature Vectors

- Energy-based Self-Supervised Learning
  - SimCLR [7]
  - MoCo [9]
  - SwAV [8]
  - BYOL [10]
  - SimSiam [11]

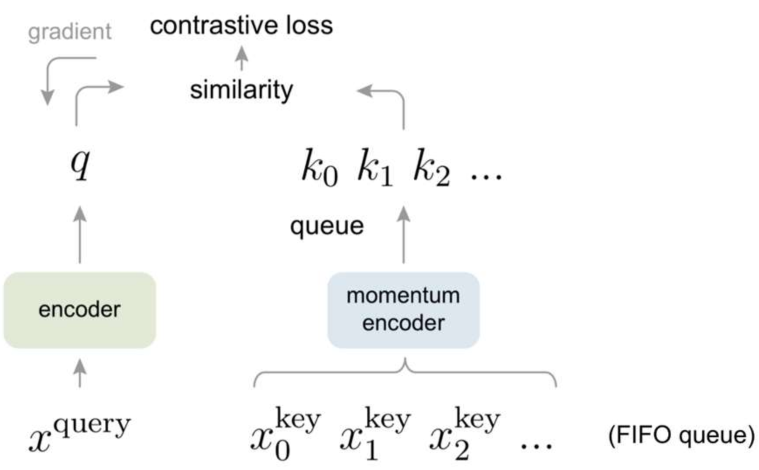


SimCLR

- Uses a large batch size (4096) to provide negatives within batch
- Requires multi-node training

MoCO

- Uses a momentum queue to store negatives
- Decouples batch size from negative set size
- Additional memory overhead, and implementation complexity

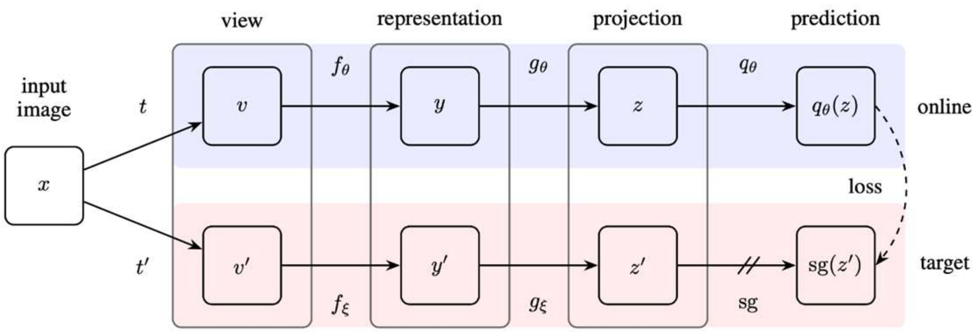


SwAV

- A cluster-center based output representation, p is used to pick center
- Key: making sure that cluster sizes are balanced (Sinkhorn-Knopp)
- Constant solution is less likely because otherwise all points are assigned to a singular cluster

BYOL

- Introduces an additional MLP (predictor), and uses momentum encoder
  - Momentum encoder
    - Exponential moving average (EMA) of base encoder weights
    - So, weights are not updated by gradients
    - But need to maintain two copies of weights
- Only uses positive samples in the loss function



Self-Supervised vs ImageNet Pretrained

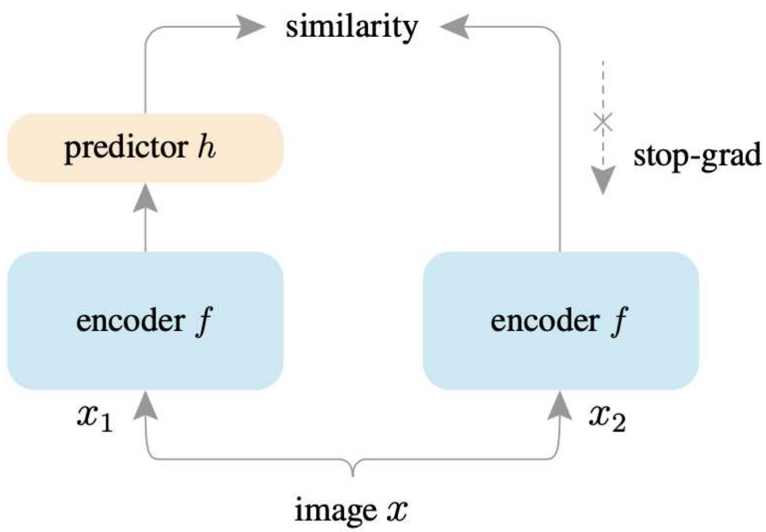
	EuroSAT	RESISC45
SwAV	96.72	92.61
MoCo v2	97.2	85.4
BYOL	97.6	88.2
SimCLR v2	97.5	85.8
ImageNet Supervised	94.98	88.56

Performance of ResNet50 on Remote Sensing Datasets

Source

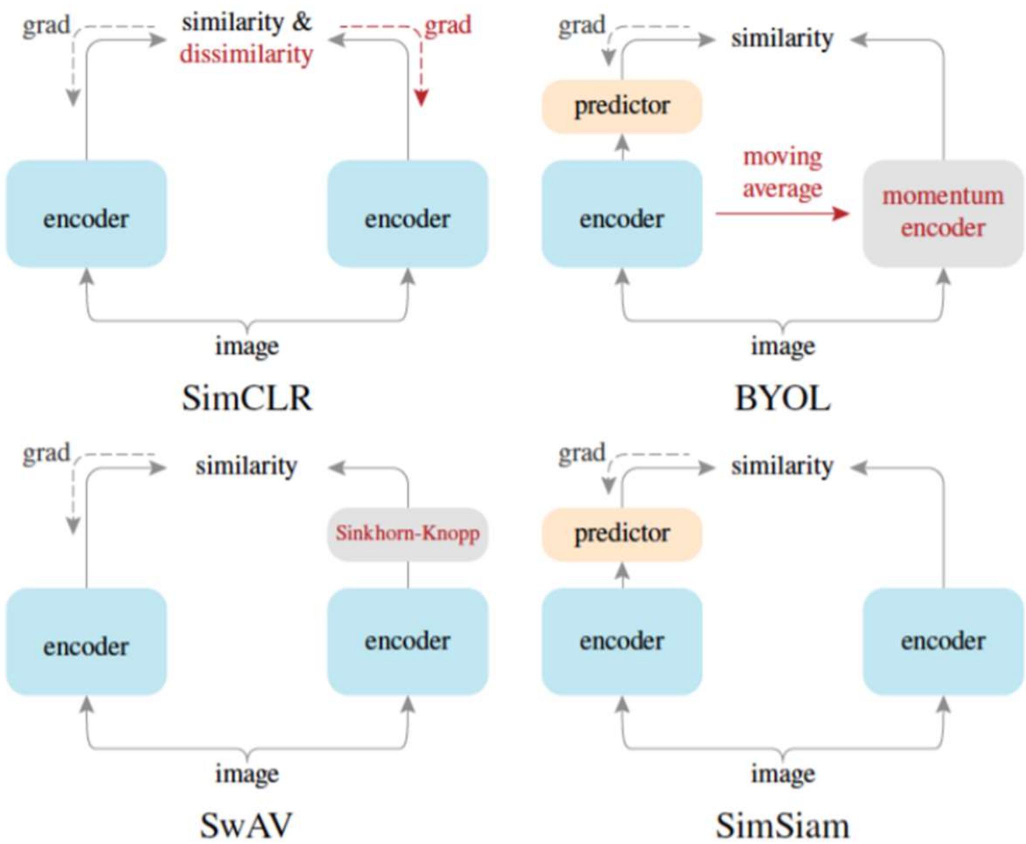
SimSiam

- SimCLR w/o negatives
- SwAV w/o online clustering
- BYOL w/o momentum encoder
- MoCo w/o negatives or momentum encoder

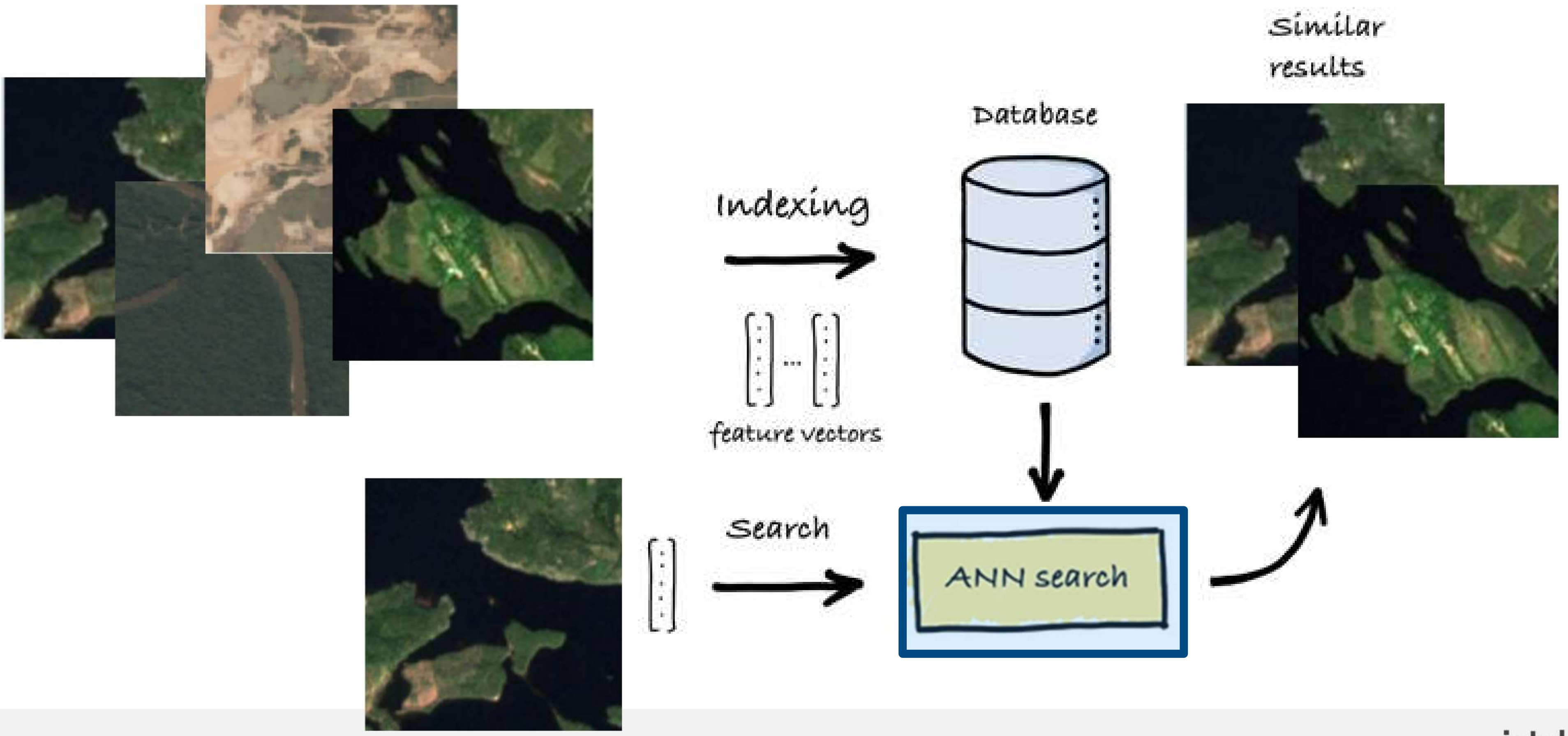


method	batch size	negative pairs	momentum encoder	100 ep	200 ep	400 ep	800 ep
SimCLR (repro.+)	4096	✓		66.5	68.3	69.8	70.4
MoCo v2 (repro.+)	**256**	✓	✓	67.4	69.9	71.0	72.2
BYOL (repro.)	4096		✓	66.5	**70.6**	**73.2**	**74.3**
SwAV (repro.+)	4096			66.5	69.1	70.7	71.8
SimSiam	**256**			**68.1**	70.0	70.8	71.3

Unifying view with SimSiam:



Approximate Nearest Neighbor (ANN) Search

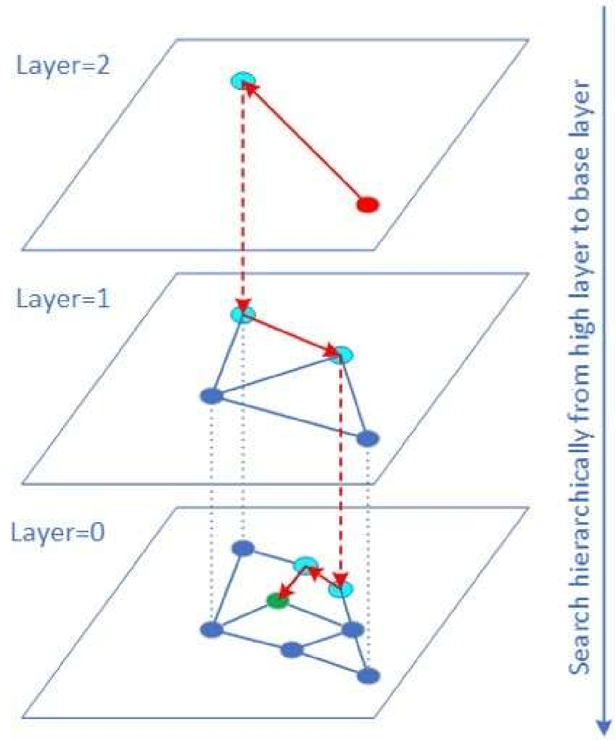


Approximate Nearest Neighbor (ANN) Search

- Tree-based space partition methods (Annoy)
- Hash based methods(Locality Sensitive Hashing)
- Clustering methods (Inverted File Index)
- Neighbourhood based methods (Hierarchical Navigable Small World Graphs, DiskANN, GraphANN)

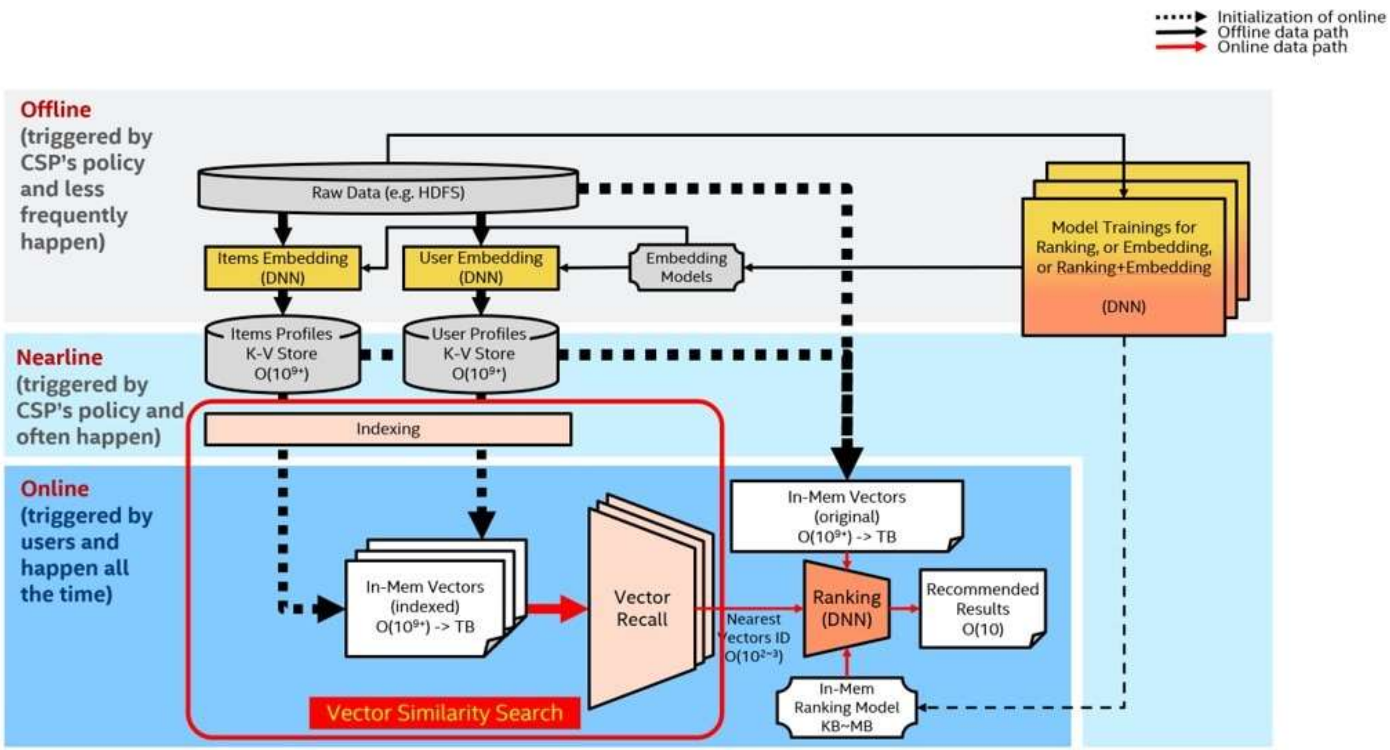
Source

Hierarchical Navigable Small World Graphs



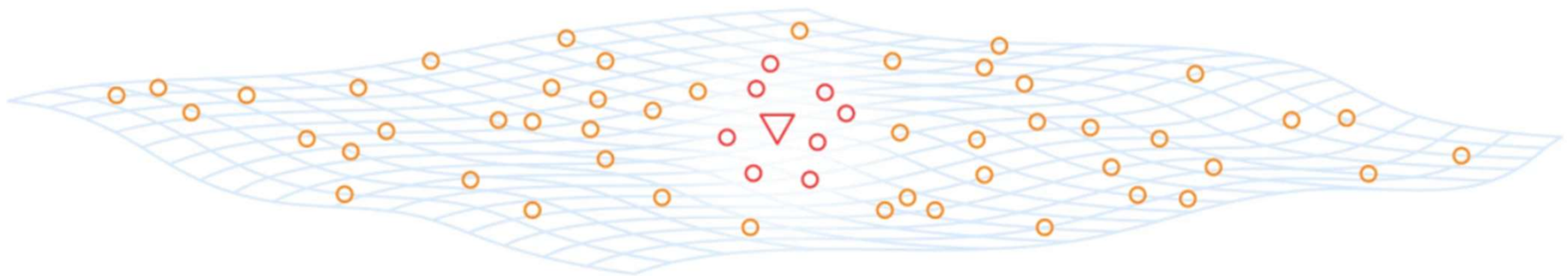
The search process through the multi-layer structure of an HNSW graph.

Accelerating Vector Recall with Intel® Deep Learning Boost VNNI and Intel® Optane



Deep1B - Dataset

- Benchmarks for Billion-Scale Similarity Search
  - Consists of 10^9 image embeddings produced as the outputs from the last fully-connected layer of the GoogLeNet model



Source

NeurIPS

- Consists of 109 image embeddings produced as the outputs from the last fully-connected layer of the GoogLeNet model

Deep1B Recall Rankings

Rank	Submission	Team	Hardware	Status	R@10
1	optanne_graphann	Intel	Intel Optane	final	0.99882
2	diskann	Microsoft Research India(*org*)	Dell PowerEdge	final	0.99821
3	cuanns_ivfpq	NVidia	NVidia GPU	final	0.99543
4	cuanns_multigpu	NVidia	NVidia GPU	final	0.99504
5	gemini	GSI Technology(*org*)	LedaE APU	final	0.99208
6	faiss_t3	Facebook Research(*org*)	NVidia GPU	final	0.94275

Source

GraphANN

- Extension of DiskANN (An extension of HNSW)
- Written in Julia
- Graph stored in Intel® Optane™ PMem
- Feature Vectors stored in DRAM
- Uses VNNI instructions for distance computation
- Static sizing of data vectors
- Optimized memory alignment and others