Machine Learning for Images

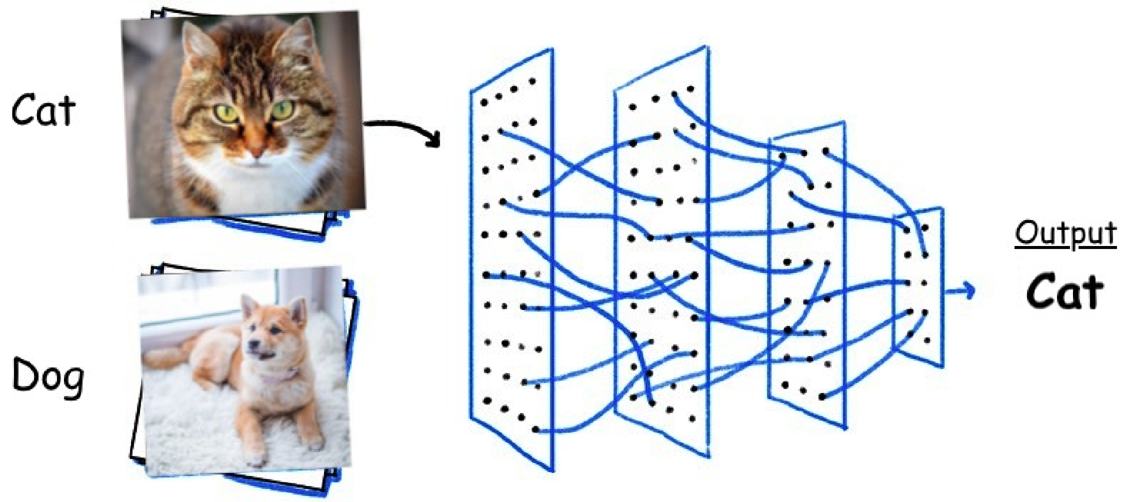
Tushar B. Kute,
http://tusharkute.com

Image Classification

- What?
  - Assigning labels to an entire image.
- Where:
  - Identifying objects in photos (e.g., cats, dogs, cars
  - Diagnosing medical images (e.g., detecting tumors in X-rays or MRIs).
- How?
  - Convolutional Neural Networks (CNNs).
  - Pre-trained models like ResNet, VGG, and Inception.

Image Classification

- Image classification is a supervised learning problem: define a set of target classes (objects to identify in images), and train a model to recognize them using labeled example photos.



Unsupervised Learning

- Unsupervised learning technique is a fully automated method that does not leverage training data.
- This means machine learning algorithms are used to analyze and cluster unlabeled datasets by discovering hidden patterns or data groups without the need for human intervention.
- With the help of a suitable algorithm, the particular characterizations of an image are recognized systematically during the image processing stage.

Supervised Classification

- Supervised image classification methods use previously classified reference samples (the ground truth) in order to train the classifier and subsequently classify new, unknown data.
- Therefore, the supervised classification technique is the process of visually choosing samples of training data within the image and allocating them to pre-chosen categories, including vegetation, roads, water resources, and buildings.
- This is done to create statistical measures to be applied to the overall image.

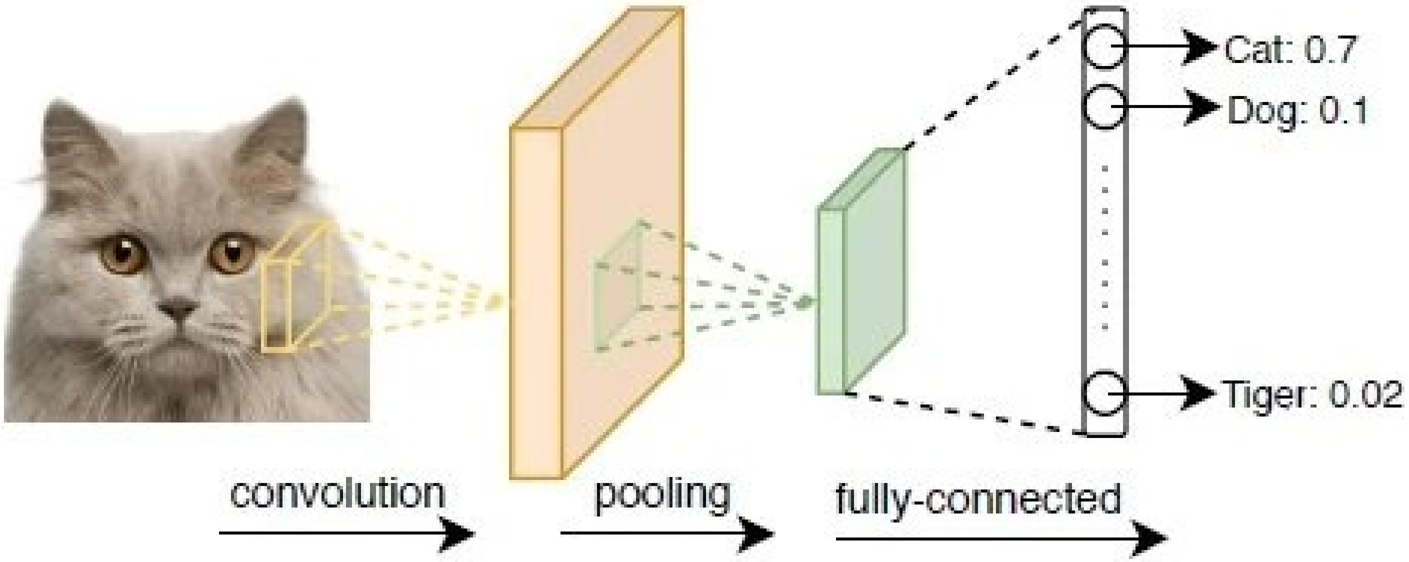
Supervised Classification

INPUT IMAGE

notebook

notebook	74%
laptop	10%
desktop computer	8%
screen	1%
space bar	1%

Convolutional Neural Network

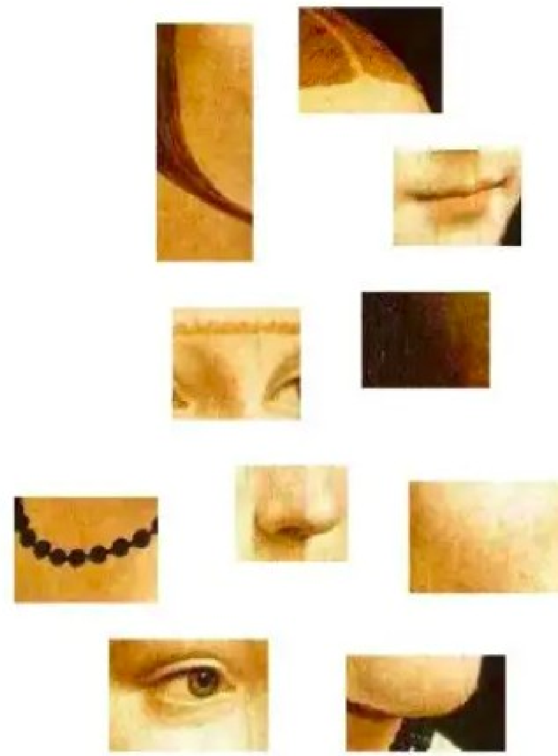


Bag of Features

- Bag of Features (BoF) is a technique used in computer vision and image processing to extract and represent features from images in a compact and meaningful way.
- The basic idea behind BoF is to extract local features from an image, such as SIFT, SURF, or ORB, and then use clustering techniques to group the features into a set of visual words.
- Each image is then represented by a histogram of these visual words, which is called a bag of features.

Bag of Features

Face



features

Bag of Features

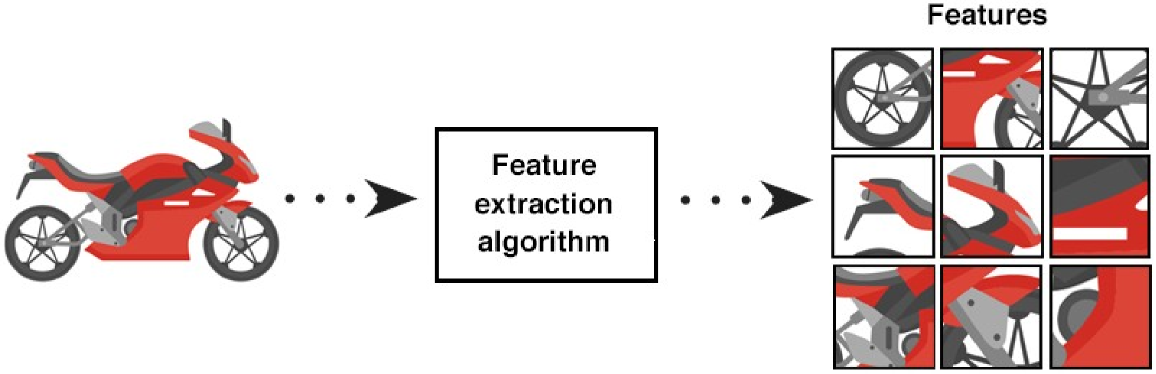
- The Bag of features representation is used in many computer vision and image processing tasks such as image retrieval, object recognition, and semantic segmentation.
- In image retrieval, BoF is used to represent images compactly and efficiently, allowing for fast and accurate retrieval of similar images.
- In object recognition, BoF extracts features from images and trains a classifier to recognize objects in new images.
- In semantic segmentation, BoF is used to extract features from images and train a model to predict the semantic labels of the pixels in the image.

Steps Involved in The BoF Process

- The basic steps involved in the Bag of Features (BoF) method include
  - feature extraction,
  - clustering, and
  - histogram representation.

Feature Extraction

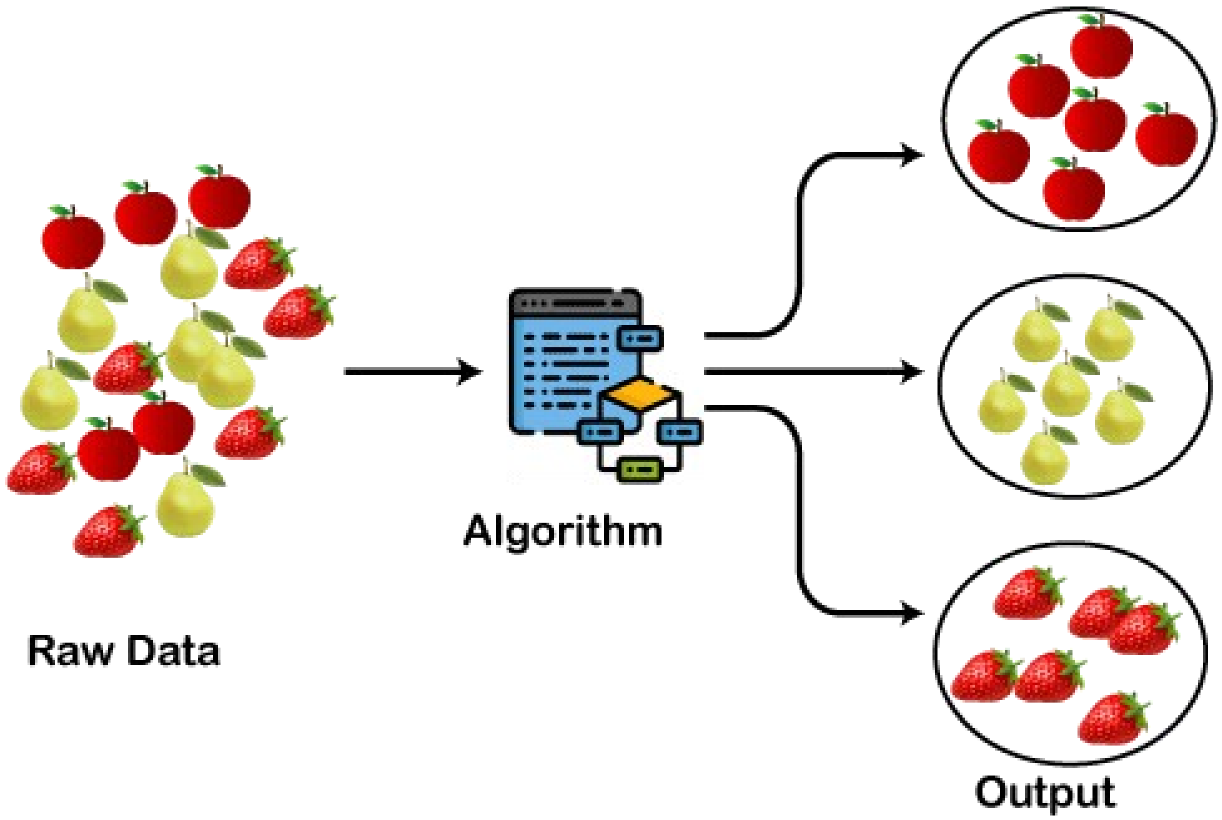
- The first step in the BoF method is to extract local features from the images.
- This is done using feature detection and description methods such as SIFT, SURF, or ORB. These methods extract a set of key points and associated descriptor vectors from the image.



Clustering

- The next step is to group the extracted features into a set of visual words.
- This is done by applying clustering techniques such as k-means or hierarchical clustering to the descriptor vectors.
- The result is a set of clusters, where each cluster represents a visual word.

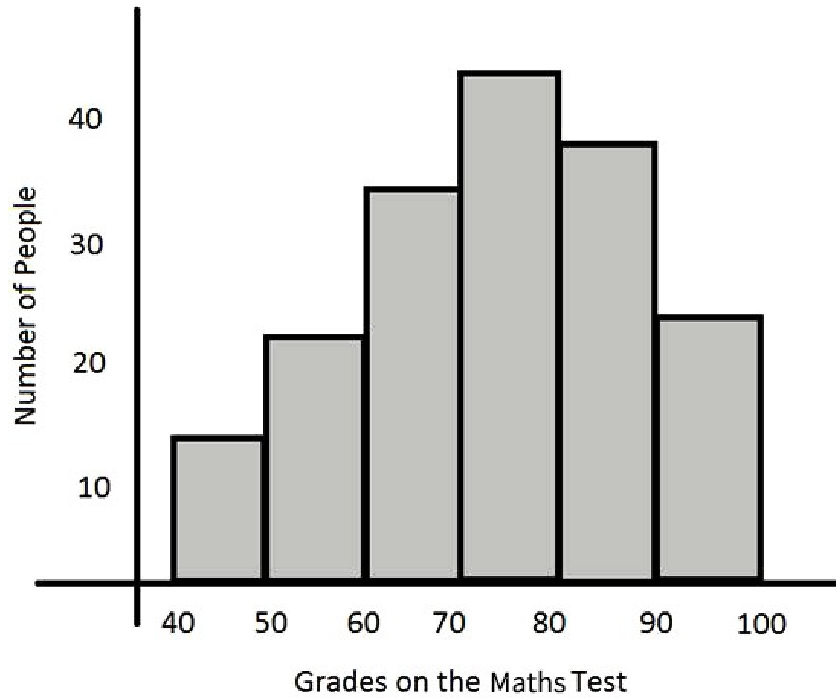
Clustering



Histogram

- Finally, the image is represented by a histogram of the visual words.
- This is done by counting the number of features that belong to each visual word and creating a histogram of these counts.
- The resulting histogram is the bag of features representation of the image.

Histogram


Number of People
40
30
20
10
40
50
60
70
80
90
100
Grades on the Maths Test

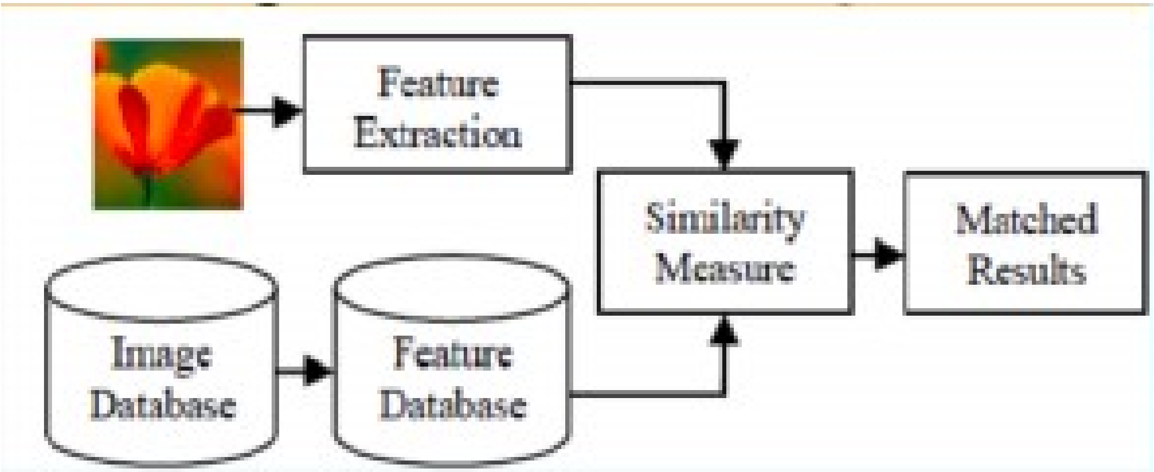
Bag of Features

- Once the feature extraction, clustering, and histogram representation steps are completed, we can use the bag of feature representation of the images to perform various tasks such as image retrieval, object recognition, and semantic segmentation.
- It is important to note that the number of clusters, or visual words, used in the BoF method will affect the representation of the images.
- Using a larger number of clusters will result in a more detailed representation, but it will also increase the computational cost and memory usage.

Bag of Features: Real World Use Cases

- Image retrieval:
  - BoF has been used in image retrieval systems to represent images compactly and efficiently, allowing for fast and accurate retrieval of similar images.
  - It has been used in applications such as image search engines, where users can search for images based on keywords or visual similarity.

Bag of Features: Real World Use Cases



Bag of Features: Real World Use Cases

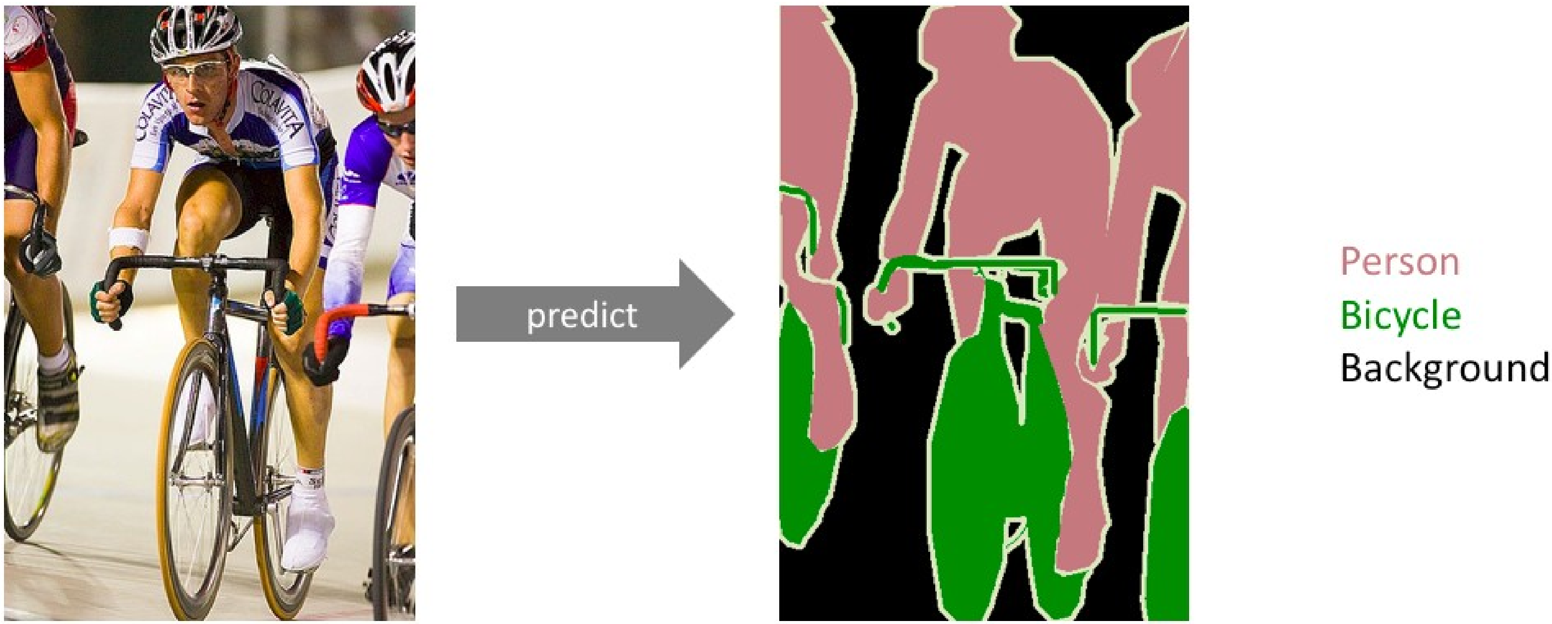
- Object recognition:
  - BoF has been used to extract features from images and train a classifier to recognize objects in new images.
  - It has been used in applications such as surveillance systems, where it can be used to recognize and track objects in video streams automatically.

Bag of Features: Real World Use Cases

Bag of Features: Real World Use Cases

- Semantic segmentation:
  - BoF has been used to extract features from images and train a model to predict the semantic labels of the pixels in the image.
  - It has been used in applications such as autonomous driving, where it can be used to segment the road, vehicles, pedestrians, and other objects in an image.

Bag of Features: Real World Use Cases



Bag of Features: Real World Use Cases

- Medical imaging:
  - BoF has been used in medical imaging to segment and classify the structures in medical images such as CT and MRI scans.
  - It allows for the automated detection and diagnosis of diseases, making it a valuable tool for radiologists and physicians.
- Robotics:
  - BoF has been used in robotics for object recognition and localization.
  - It allows robots to understand their environment and identify objects, allowing them to interact with the environment more effectively.

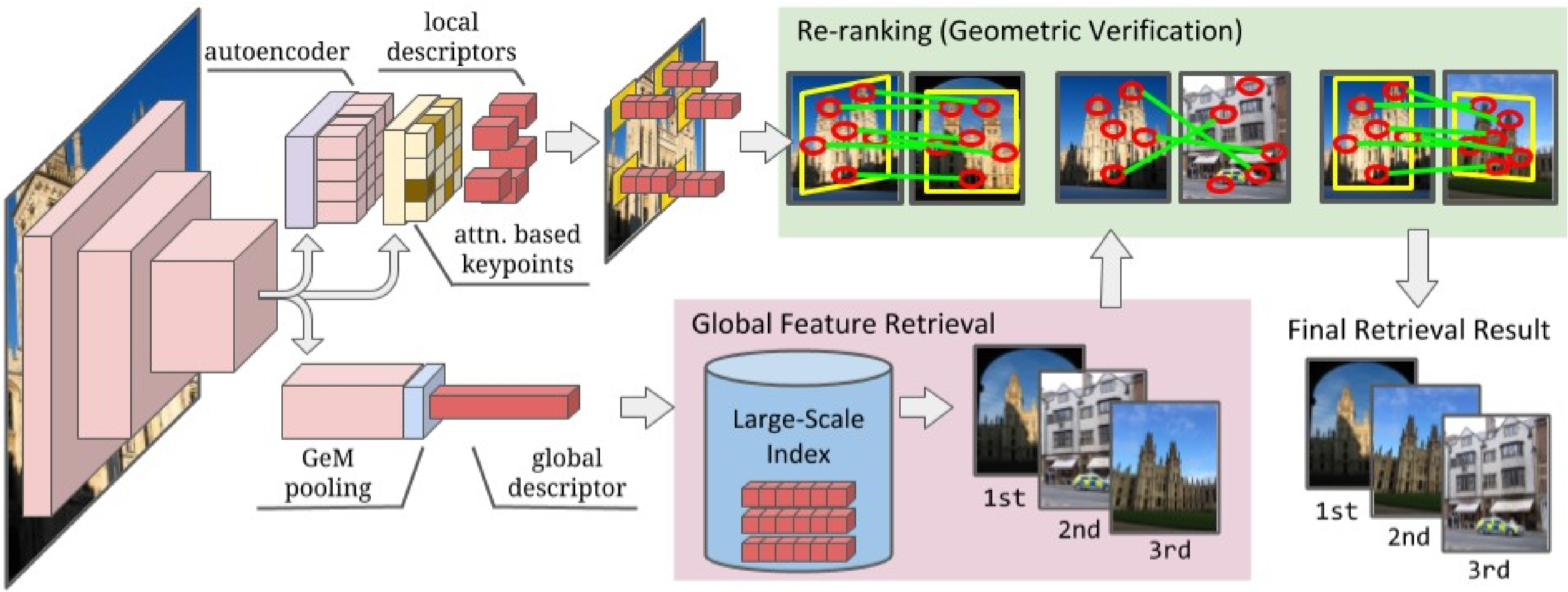
Large-scale Instance Recognition

- Large-scale instance recognition (LIR) is a computer vision task that aims to recognize a specific instance of an object in a large dataset of images.
- This is a challenging task because it requires the model to be able to identify the same object even when it appears in different poses, sizes, and illumination conditions.

Large-scale Instance Recognition

- LIR has a wide range of applications, such as:
  - Image retrieval: LIR can be used to retrieve images of a specific object from a large dataset. For example, you could use LIR to find all the images of the Eiffel Tower in a collection of millions of images.
  - Object tracking: LIR can be used to track the movement of an object in a video. This can be useful for applications such as surveillance or robotics.
  - Scene understanding: LIR can be used to understand the context of an image. For example, you could use LIR to determine that an image of a car is in a parking lot.

Large-scale Instance Recognition



Large-scale Instance Recognition

- There are a number of challenges that need to be addressed in order to achieve good performance in LIR. These challenges include:
  - Object variability: The same object can appear in a wide variety of poses, sizes, and illumination conditions. The model needs to be able to handle this variability in order to recognize the object correctly.

Large-scale Instance Recognition

- Data scarcity:
  - There is often not enough data available to train a model for LIR.
  - This is because it is difficult to collect a large dataset of images of the same object in different poses, sizes, and illumination conditions.
- Computational complexity:
  - LIR is a computationally expensive task.
  - This is because the model needs to be able to process a large number of images in order to find the instances of the target object.

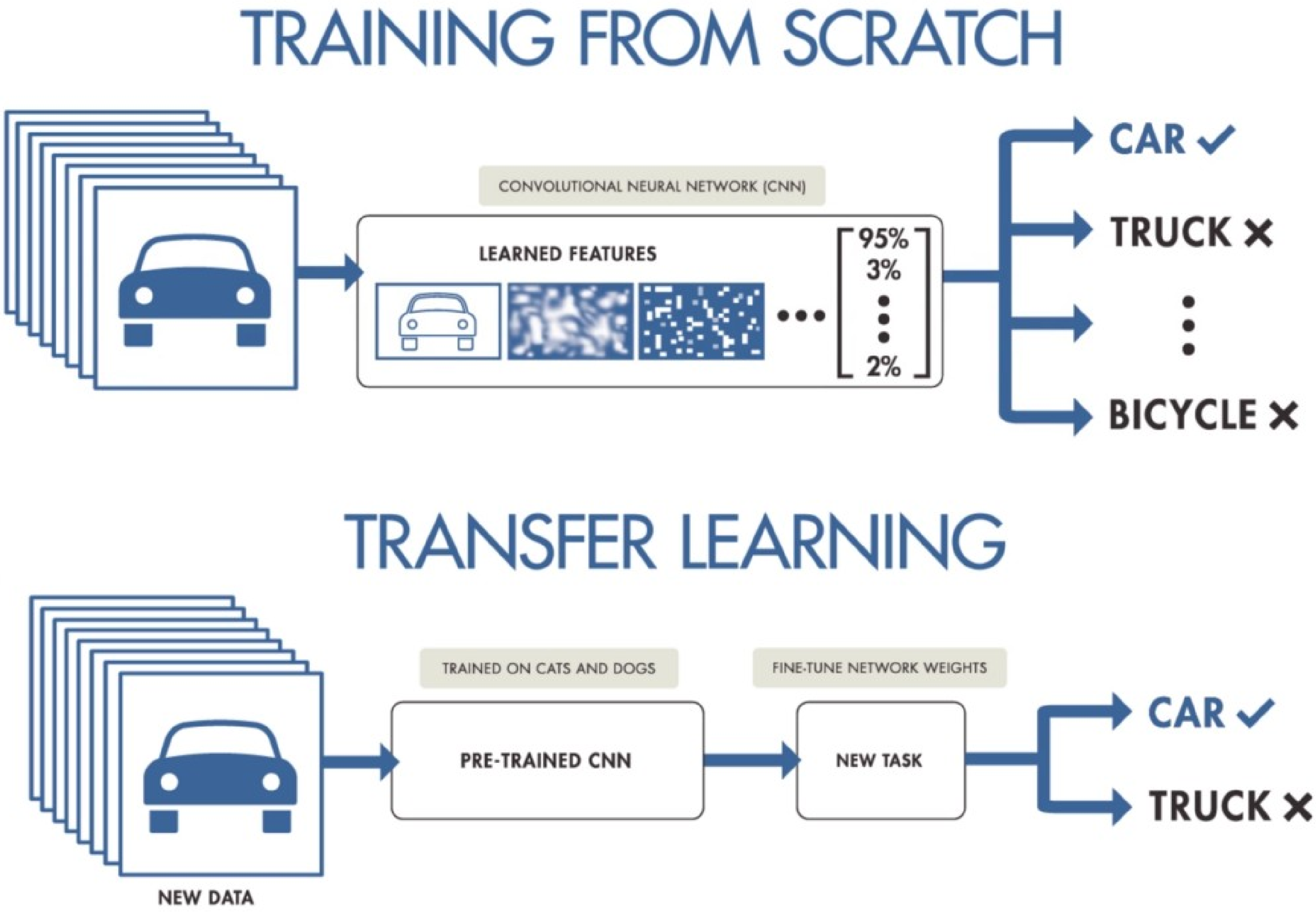
Transfer Learning

- Transfer learning is a machine learning technique where a model trained on one task is reused as the starting point for a model on a second task.
- This can be done by freezing the weights of the first model and then training the second model on a new dataset.
- This can help to improve the performance of the second model, as it can start from a point where the first model has already learned some important features.

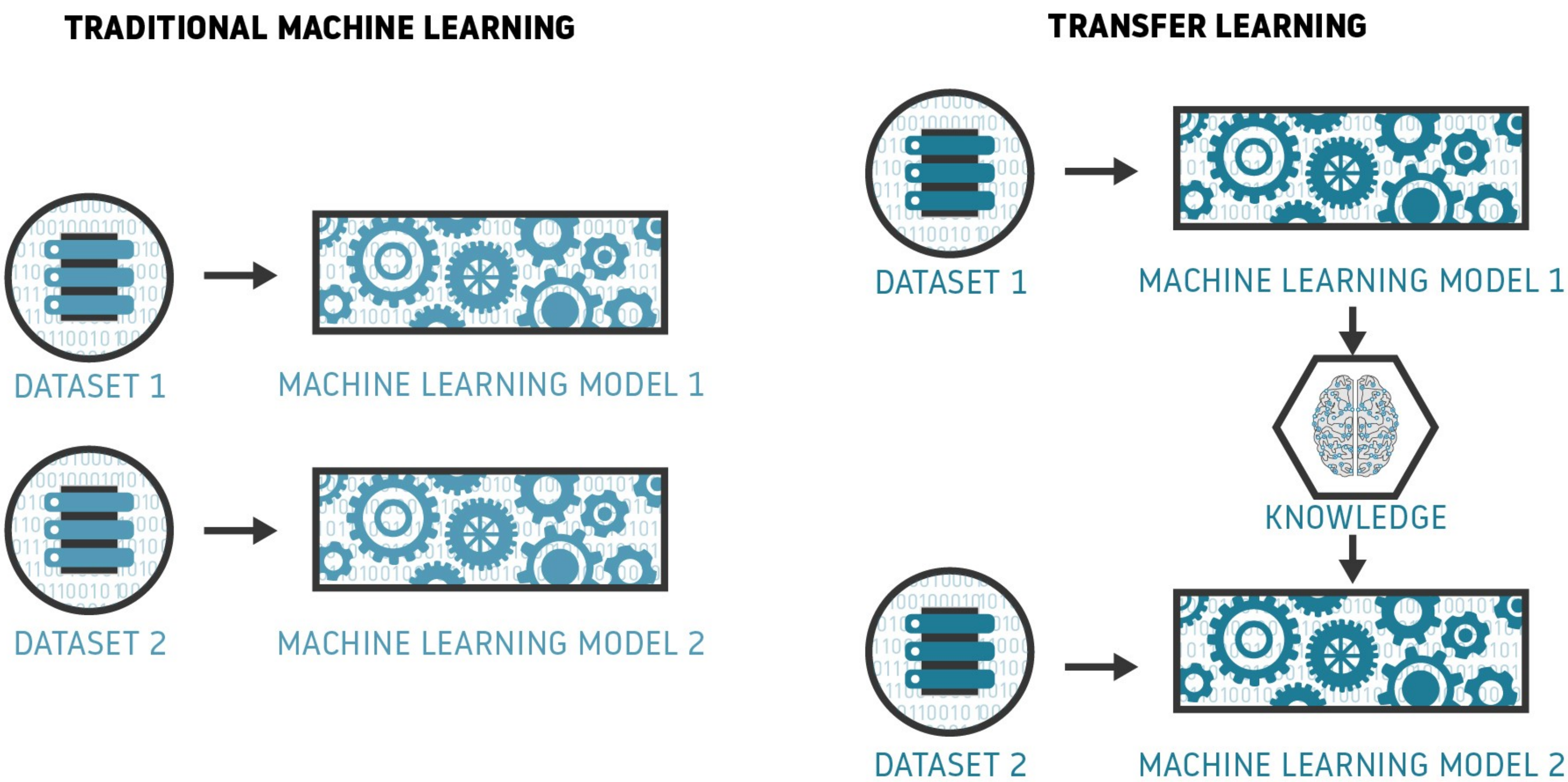
Transfer Learning

- Transfer learning is a popular approach in deep learning, as it can help to reduce the amount of data that needs to be collected for the second task.
- This can be especially useful when the second task has a small dataset, as it can be difficult to train a model from scratch on such a small dataset.

Transfer Learning



Transfer Learning



tusharkute.com

Transfer Learning: Benefits

- Reduces the amount of data needed: Transfer learning can help to reduce the amount of data needed to train a model. This can be especially useful when the second task has a small dataset.
- Improves performance: Transfer learning can help to improve the performance of a model. This is because the first model has already learned some important features, which can be reused by the second model.
- Saves time: Transfer learning can save time, as it does not require the second model to be trained from scratch.

Transfer Learning: Limitations

- The first model must be trained on a related task: The first model must be trained on a task that is related to the second task. This is because the first model will learn features that are relevant to the second task.
- The second model may not generalize well: The second model may not generalize well to new data. This is because the second model is only trained on a subset of the possible data.
- The first model may not be available: The first model may not be available. This is because the first model may have been trained by someone else, or it may not be publicly available.

Transfer Learning: Develop Model Approach

- 1. Select Source Task. You must select a related predictive modeling problem with an abundance of data where there is some relationship in the input data, output data, and/or concepts learned during the mapping from input to output data.
- 2. Develop Source Model. Next, you must develop a skillful model for this first task. The model must be better than a naive model to ensure that some feature learning has been performed.

Transfer Learning: Develop Model Approach

- 3. Reuse Model. The model fit on the source task can then be used as the starting point for a model on the second task of interest. This may involve using all or parts of the model, depending on the modeling technique used.
- 4. Tune Model. Optionally, the model may need to be adapted or refined on the input-output pair data available for the task of interest.

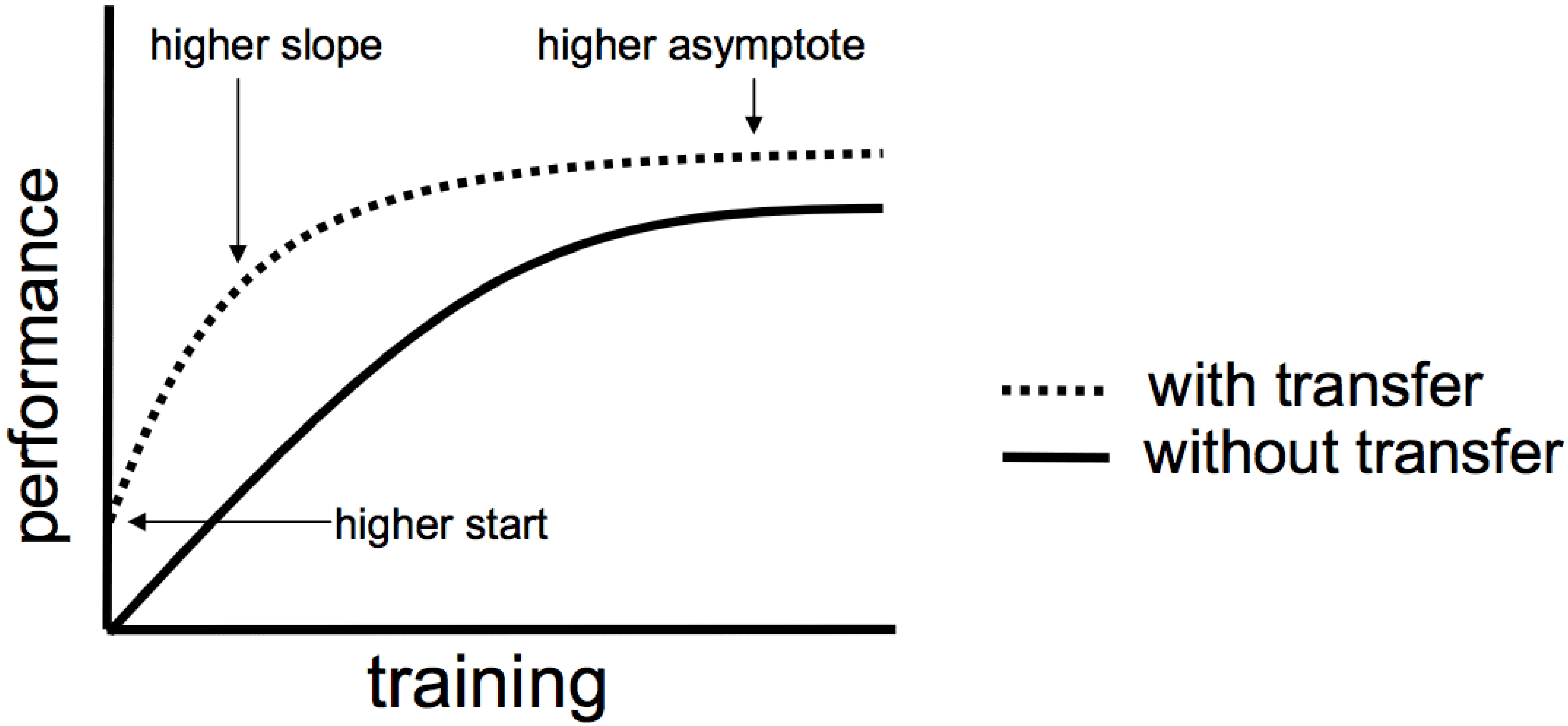
Pre-trained Model Approach

- Select Source Model. A pre-trained source model is chosen from available models. Many research institutions release models on large and challenging datasets that may be included in the pool of candidate models from which to choose from.
- Reuse Model. The model pre-trained model can then be used as the starting point for a model on the second task of interest. This may involve using all or parts of the model, depending on the modeling technique used.
- Tune Model. Optionally, the model may need to be adapted or refined on the input-output pair data available for the task of interest.

Transfer Learning: Why?

- Higher start. The initial skill (before refining the model) on the source model is higher than it otherwise would be.
- Higher slope. The rate of improvement of skill during training of the source model is steeper than it otherwise would be.
- Higher asymptote. The converged skill of the trained model is better than it otherwise would be.

Transfer Learning: Why?



Thank you

@mitu_skillologies

@mITuSkillologies

@mitu_group

@mitu-skillologies

@MITUSkillologies

kaggle

@mituskillologies

Web Resources
https://mitu.co.in
http://tusharkute.com

@mituskillologies

contact@mitu.co.in

tushar@tusharkute.com