Unsupervised ML for Classification Problems

John Branch Bedoya
Pedro Atencio Ortiz

December 01, 2020

UNIVERSIDAD NACIONAL DE COLOMBIA

Institución Universitaria ITM
Reacreditada en Alta Calidad

pedroatencio@itm.edu.co

Agenda

Session 2: Unsupervised Image Classification

1. Image classification: fundamentals.
   1. Supervised learning.
2. The unsupervised representation learning problem.
   1. Self-supervised learning. (pretext task).
3. Clustering for unlabeled classification.
   1. Simple loss clustering (MLP).
4. SCAN: Learning to Classify Images Without Labels.
   1. k-Nearest Neighbors: A prior approach.
   2. Clustering as a loss function.
5. Python practice: Image Classification.
   1. Mini-batch k-means.
   2. k-NN for nearest sample extraction.
6. Conclusions.

1. Image clasificación: Fundamentals

Classification pipeline

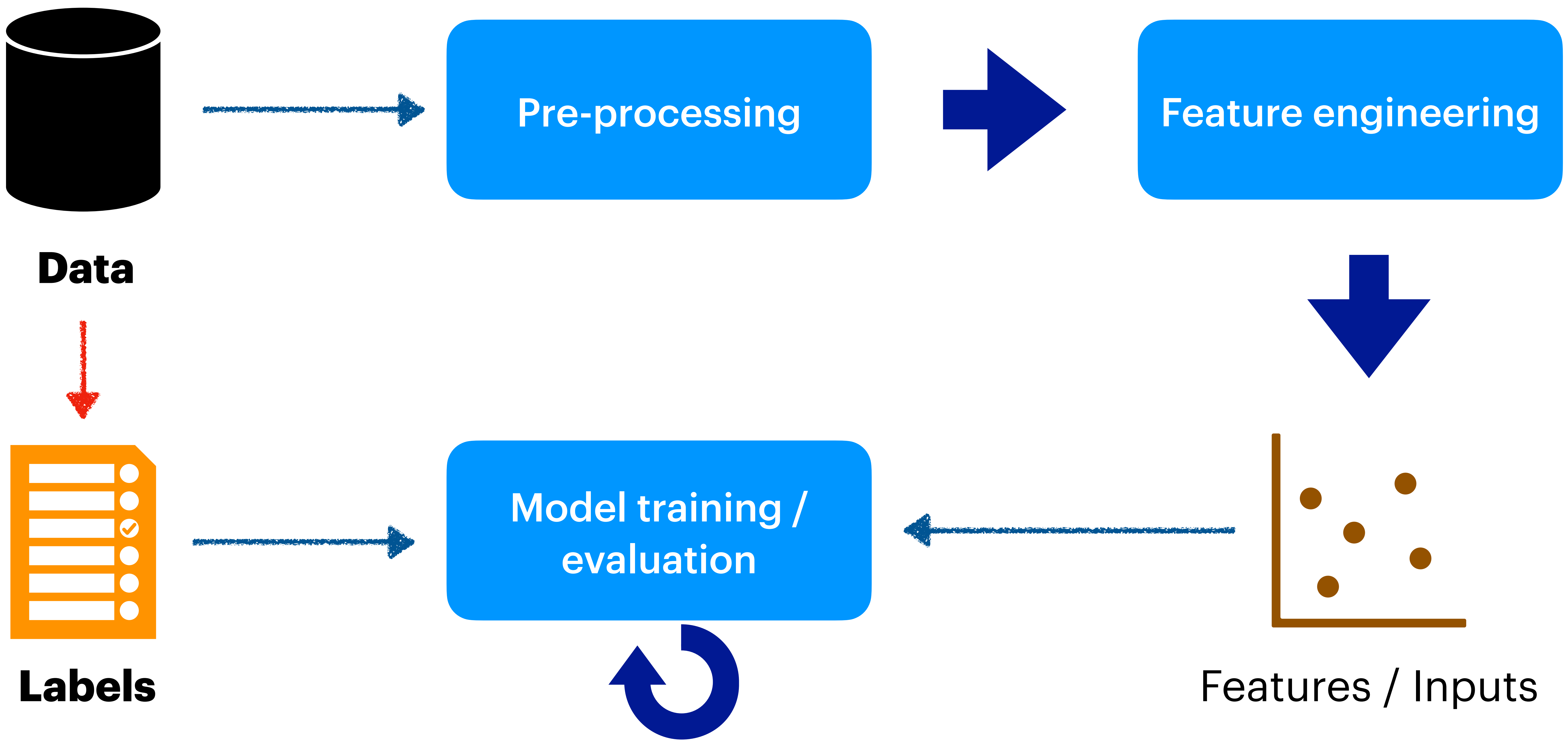


Image classification pipeline

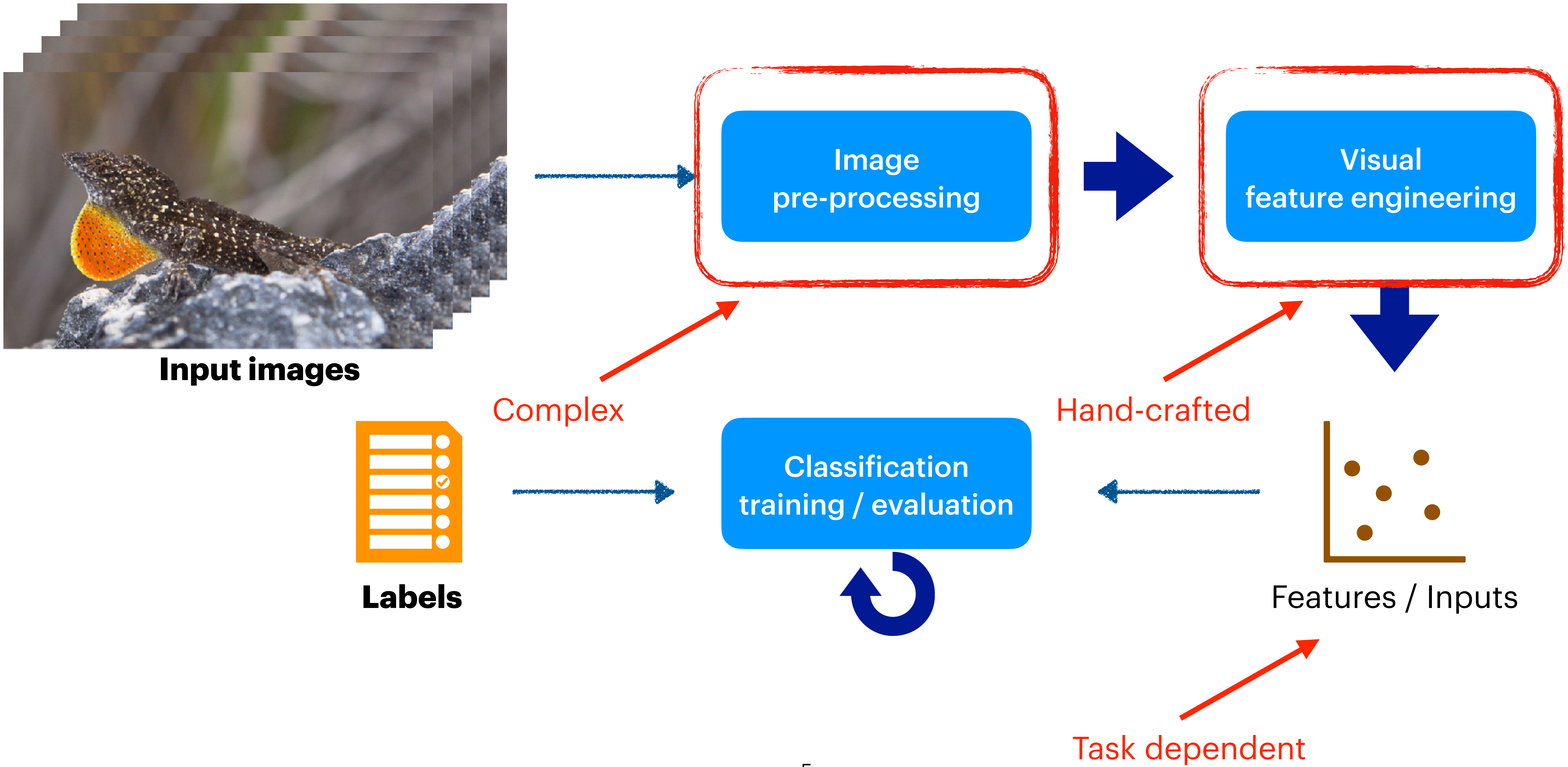
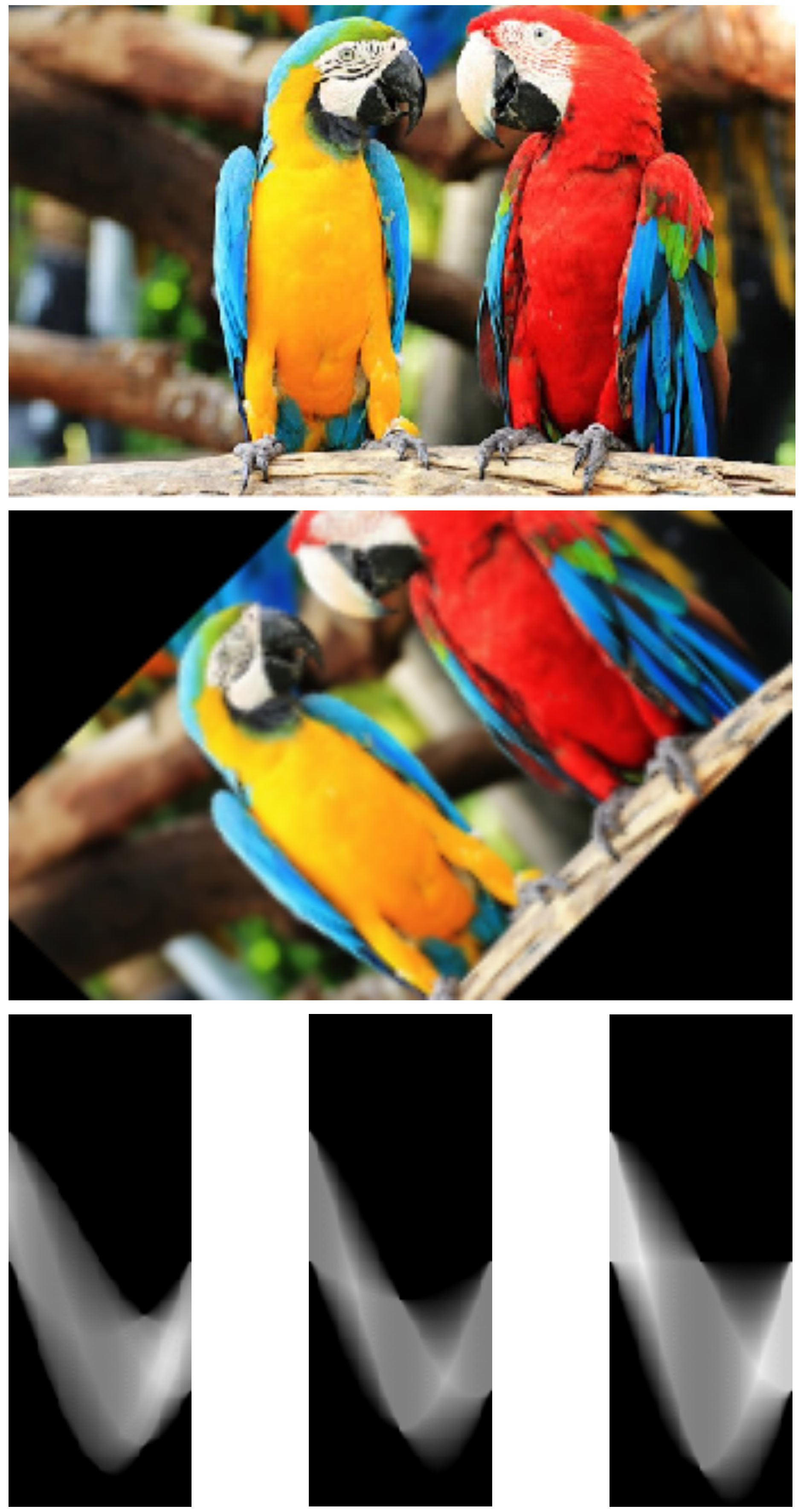


Image pre-processing

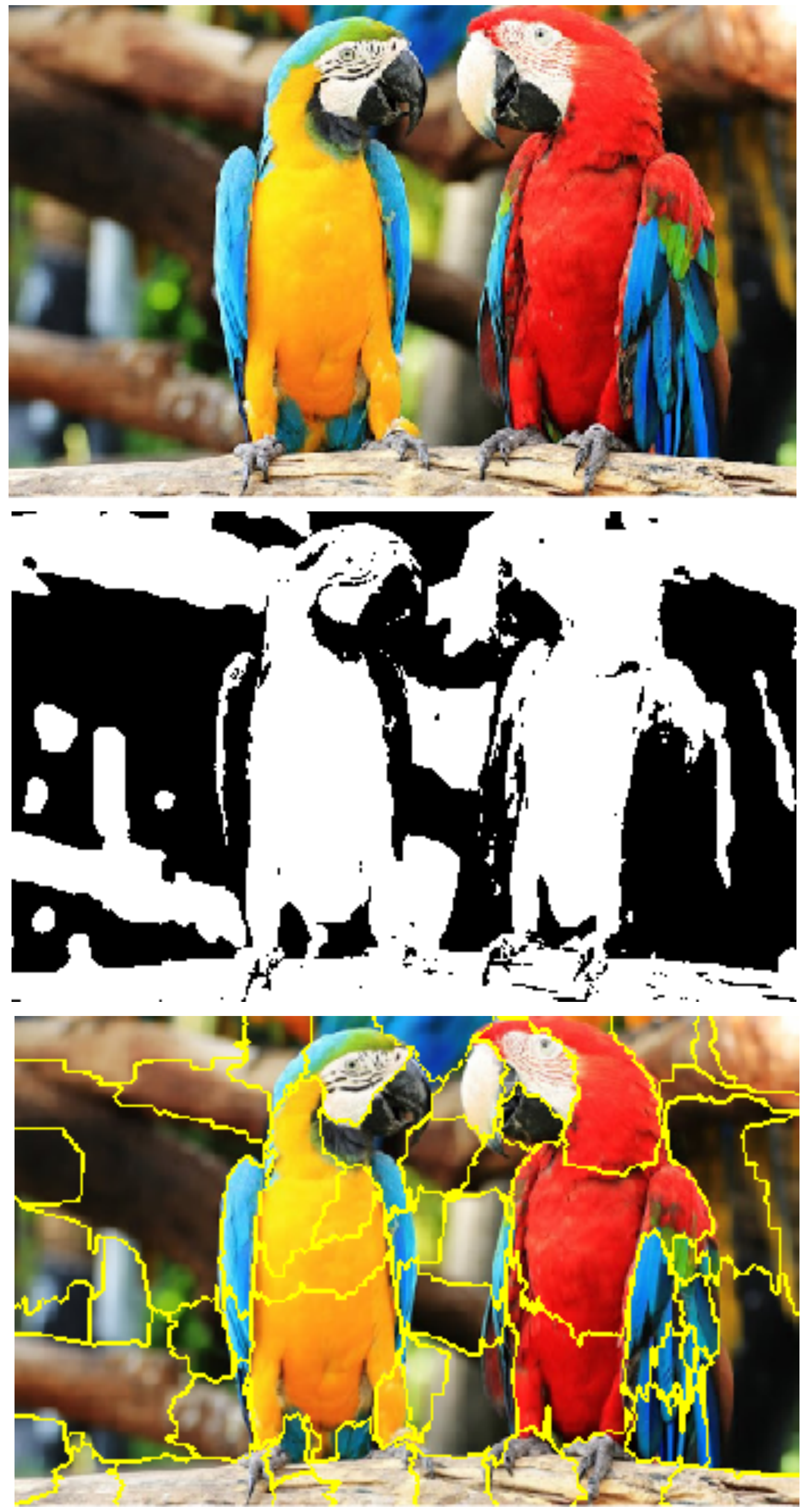
Colorimetry

Spatial filtering

Transform filtering

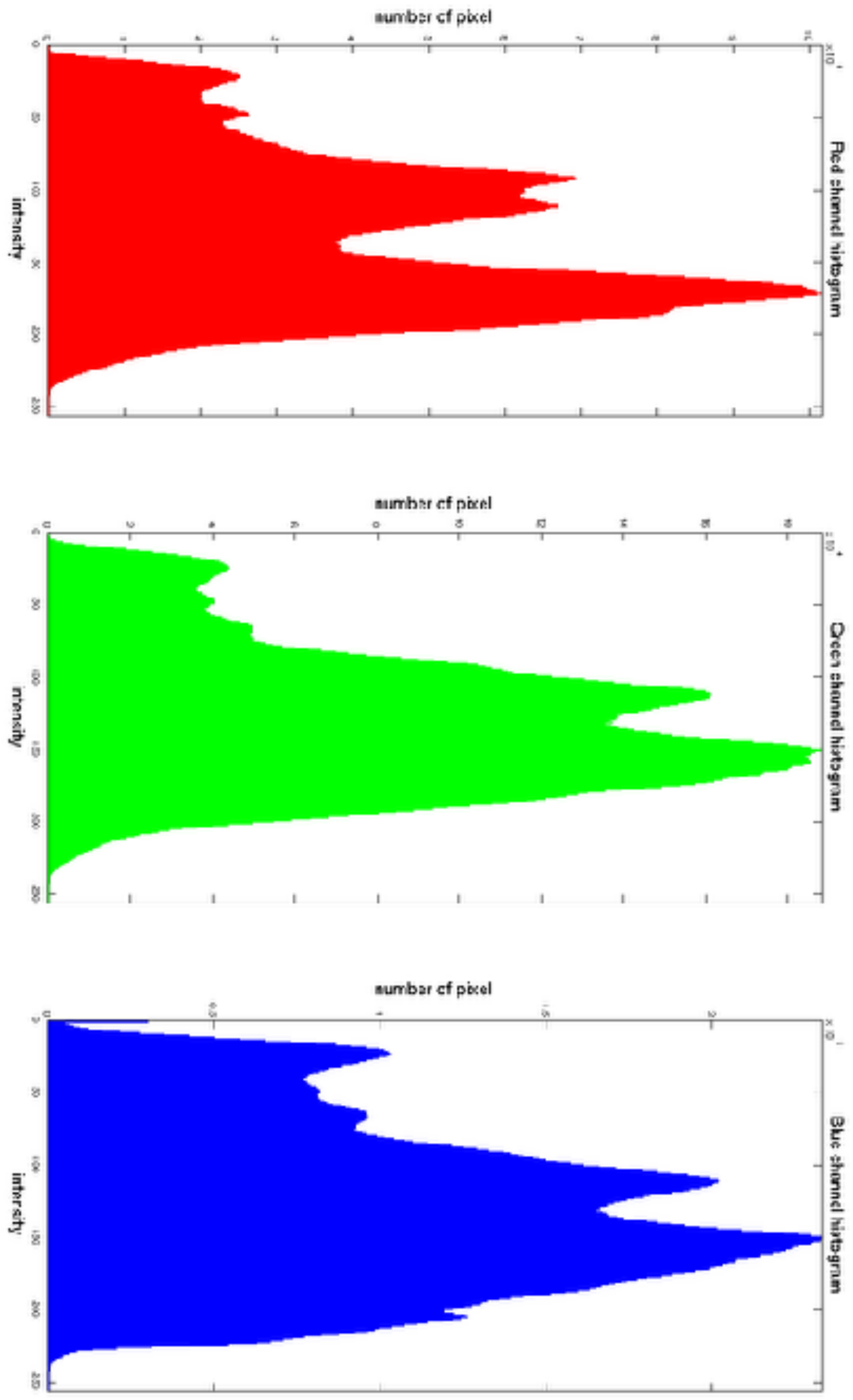


Segmentation

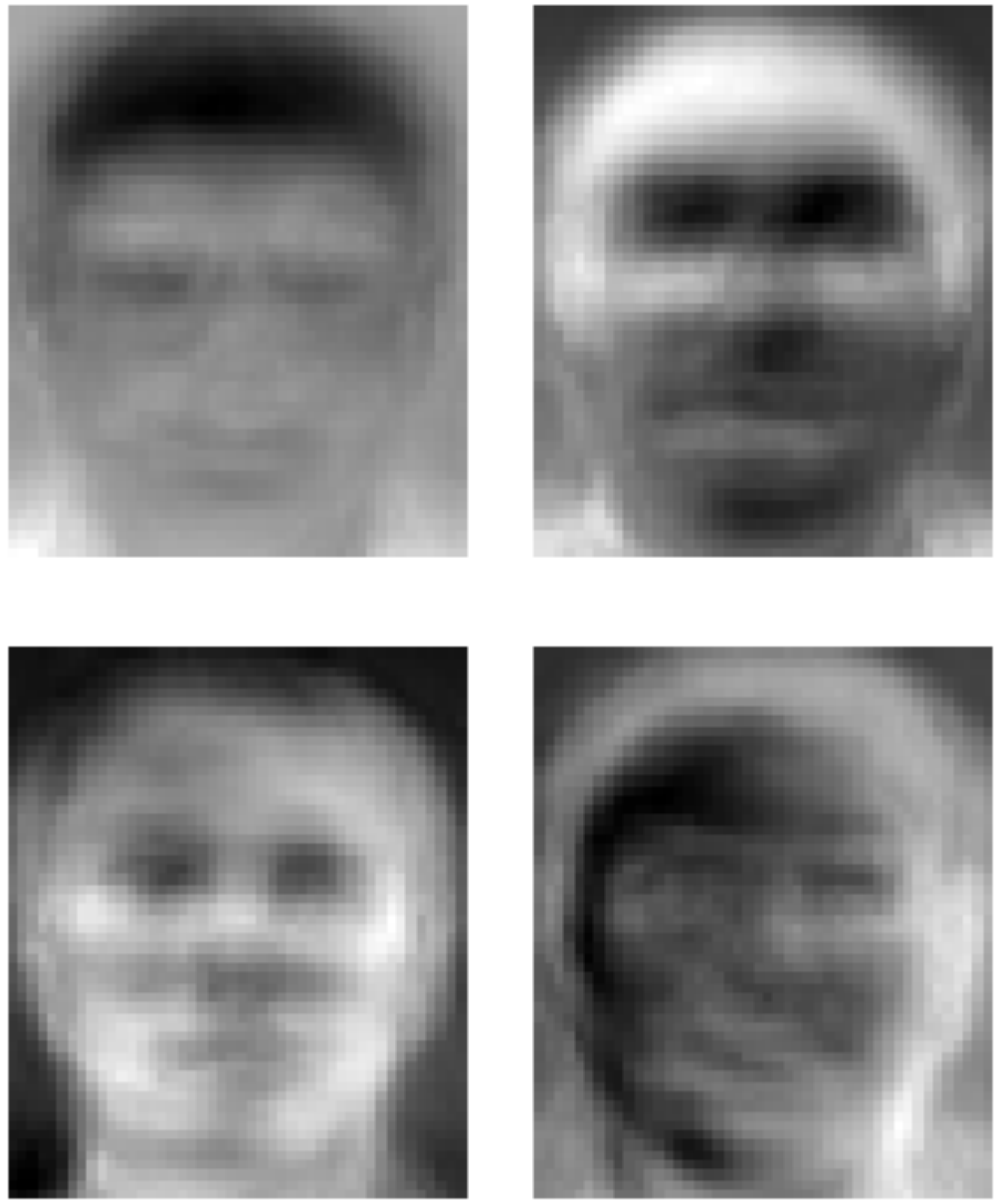


Visual feature engineering

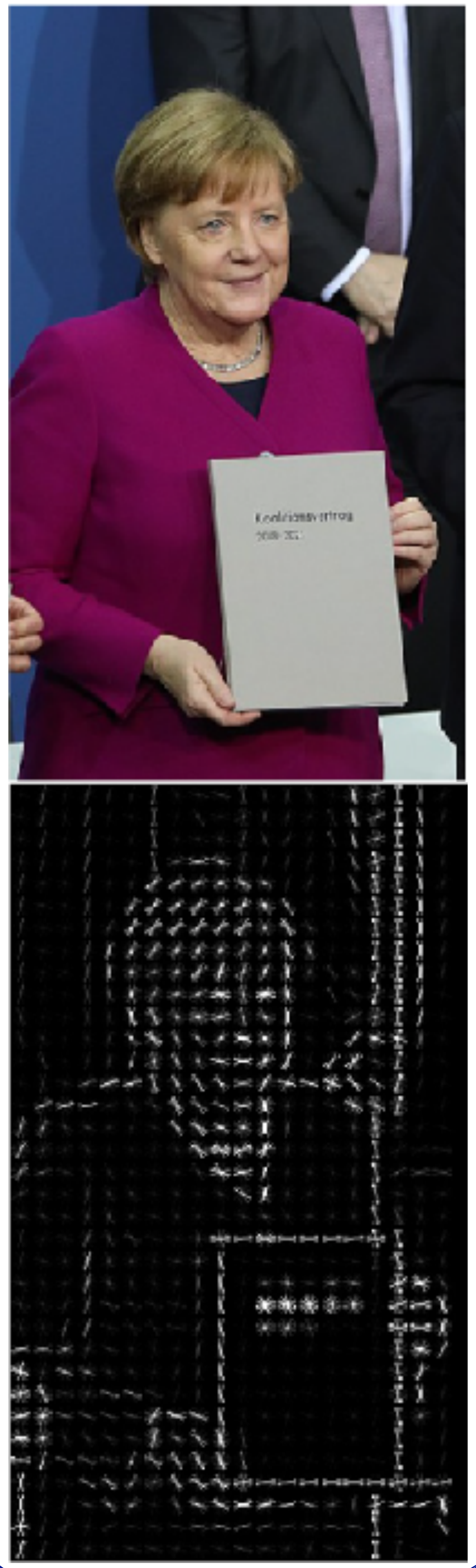
RGB Histograms



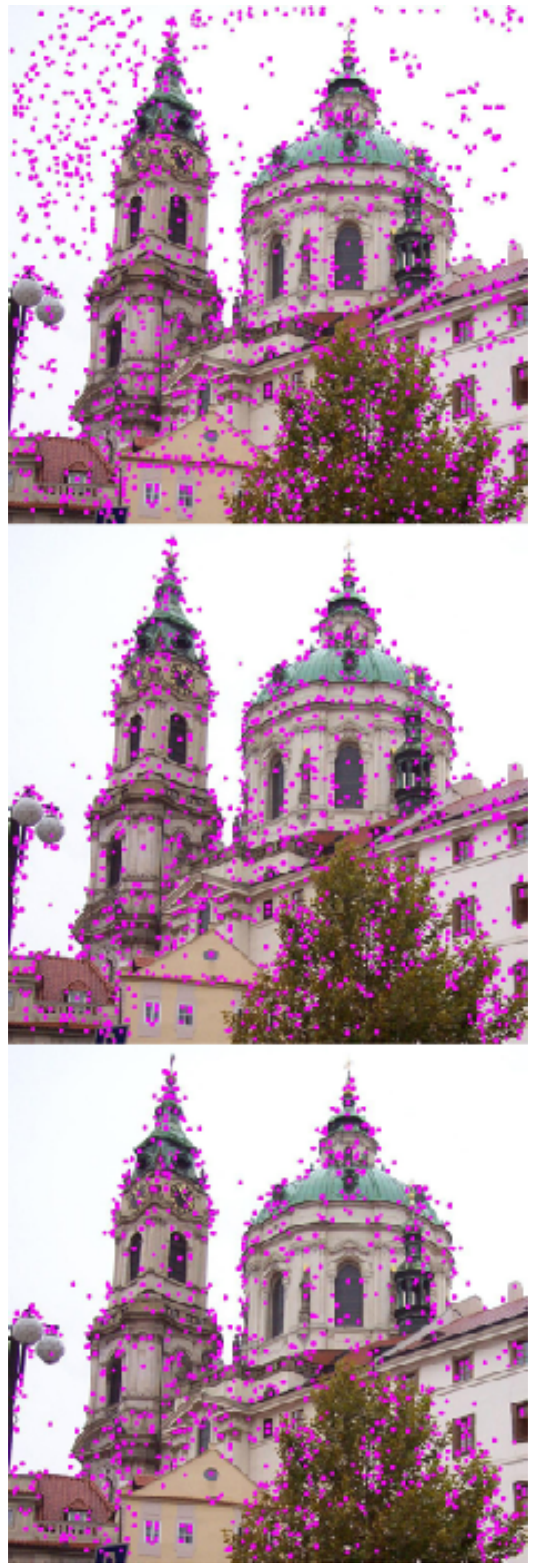
EigenFaces



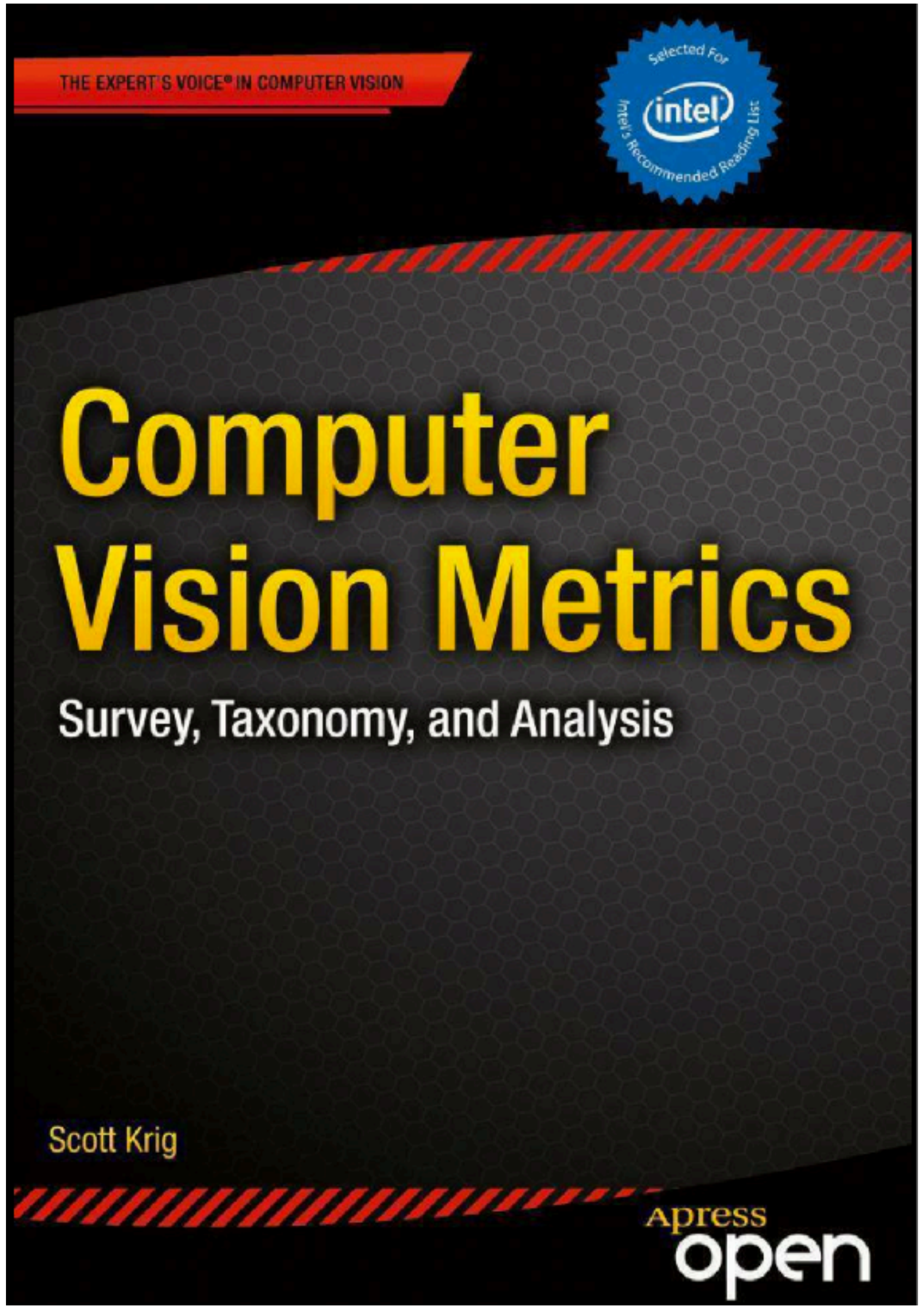
HoG



SIFT

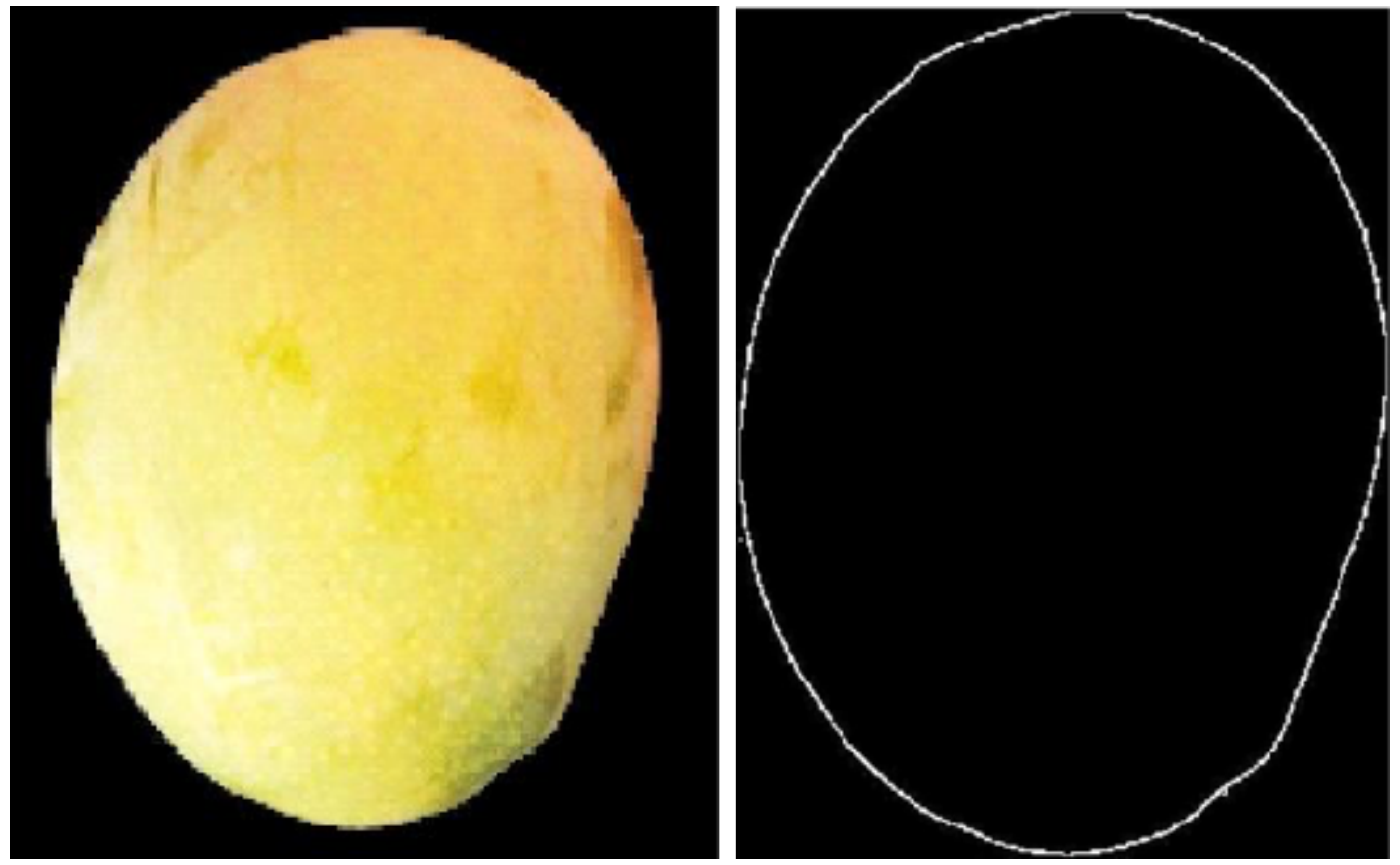


Visual feature engineering

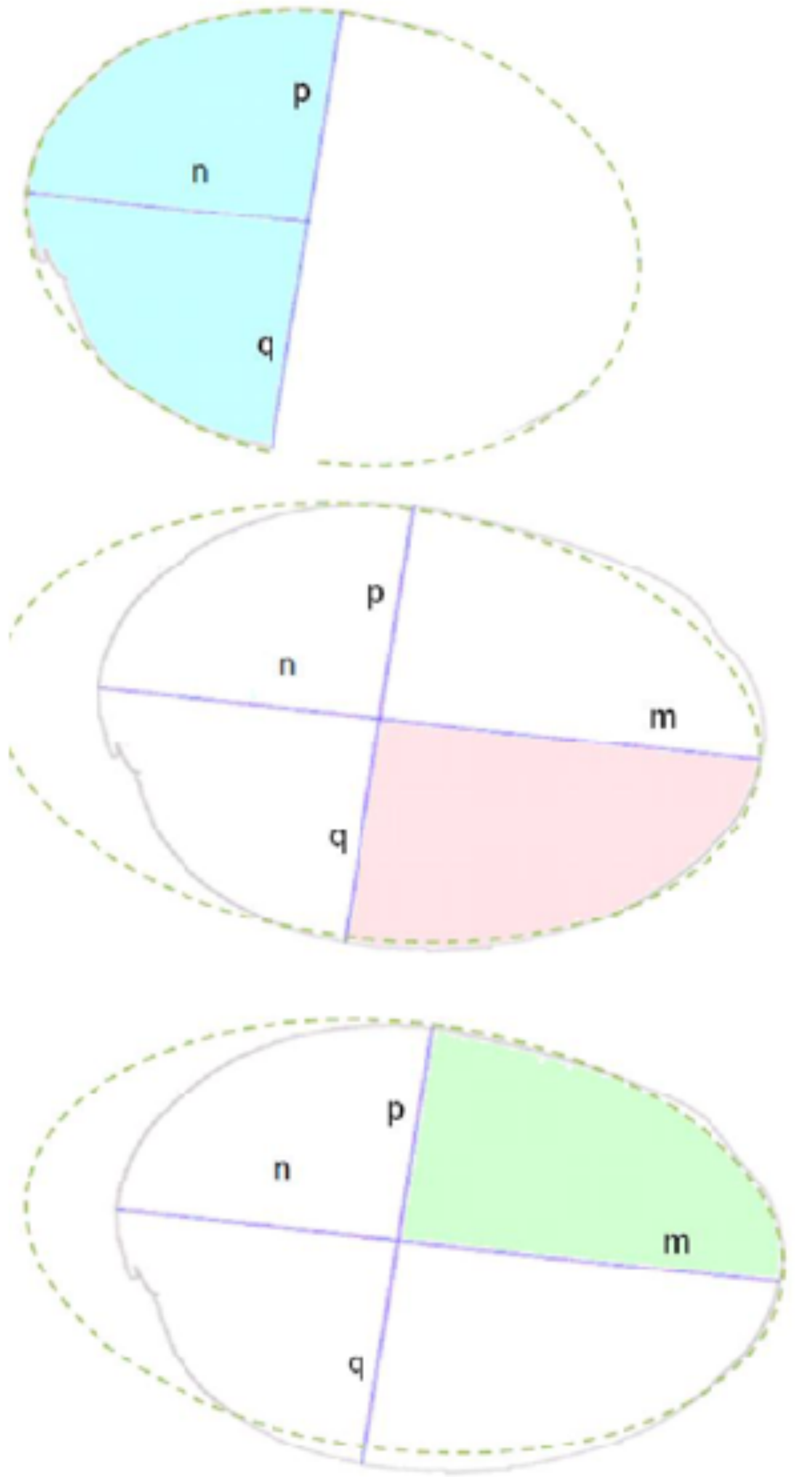


Visual feature engineering

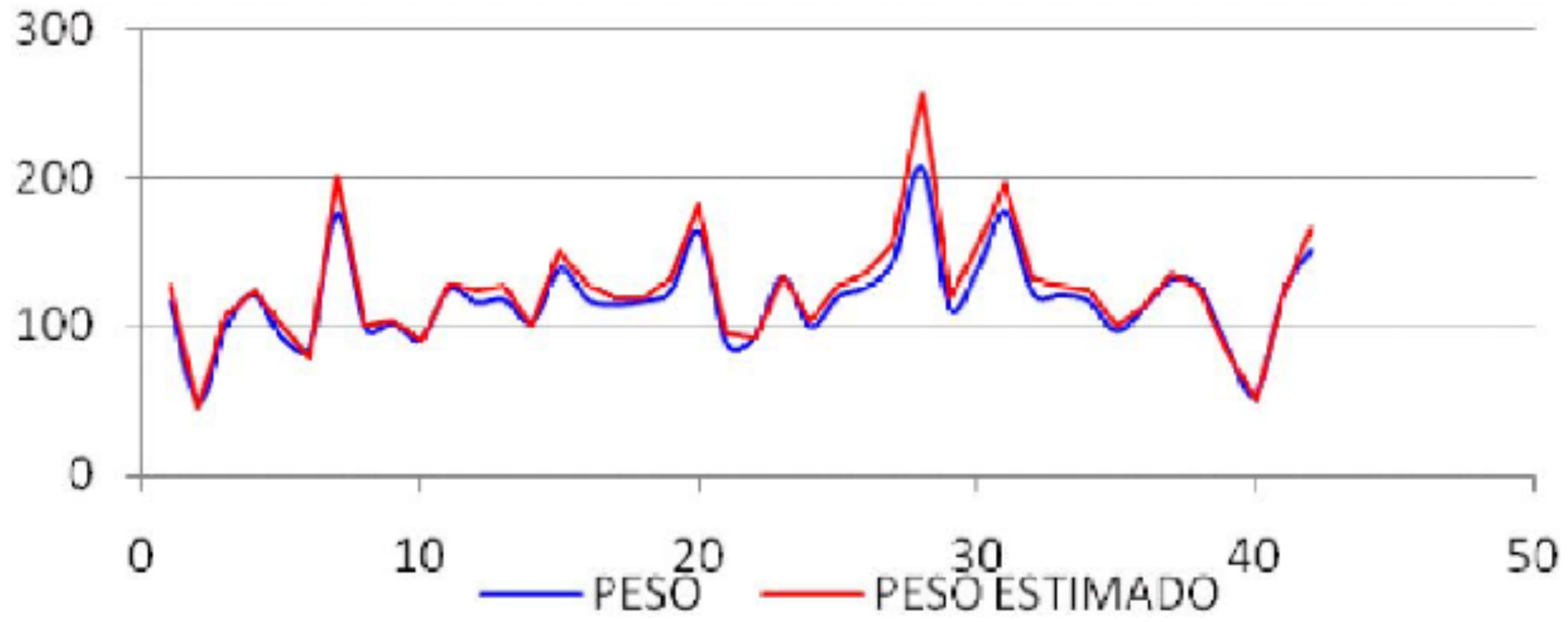
Acquisition & processing



PCA (axis detection)



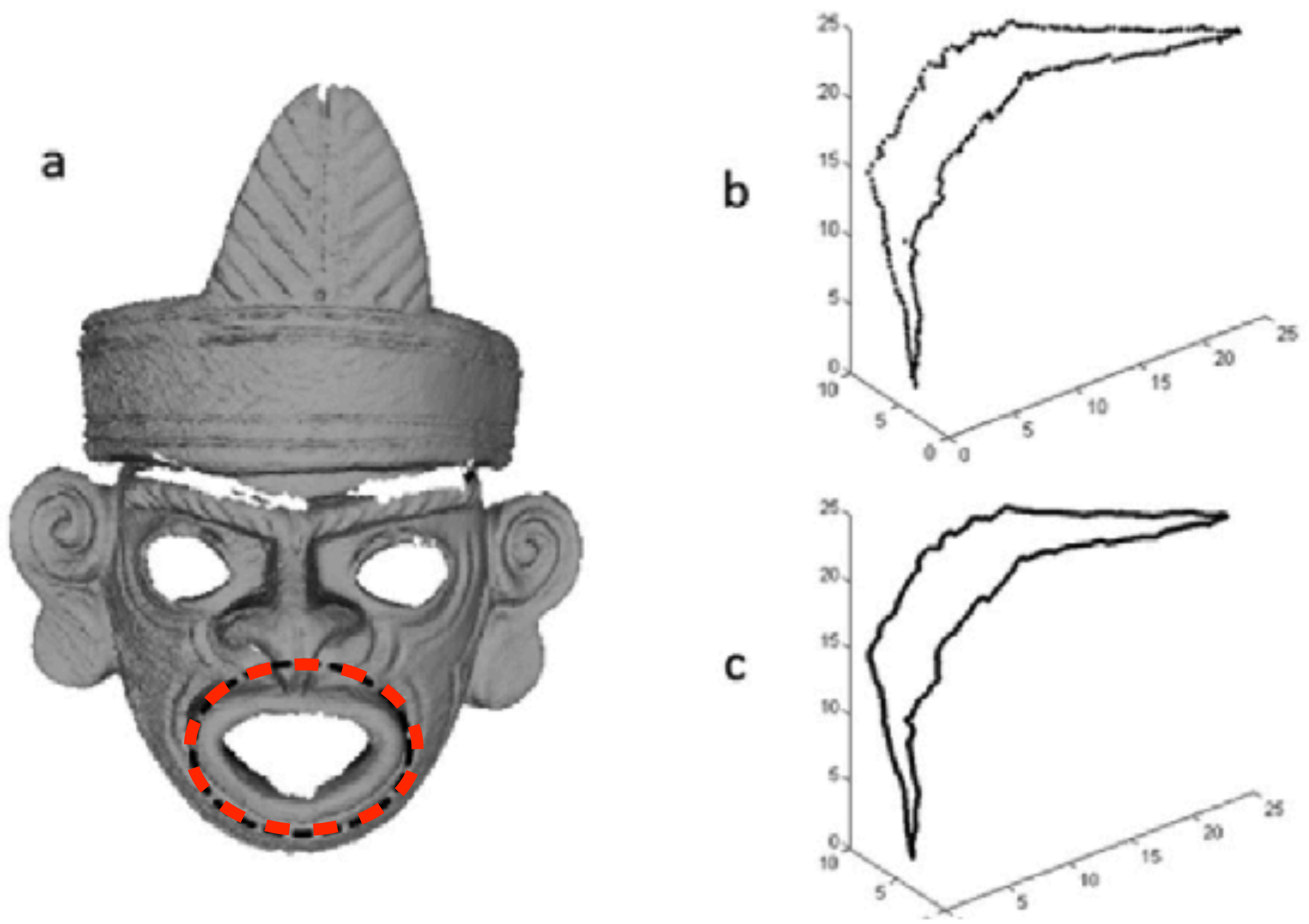
Volume / Weight regression



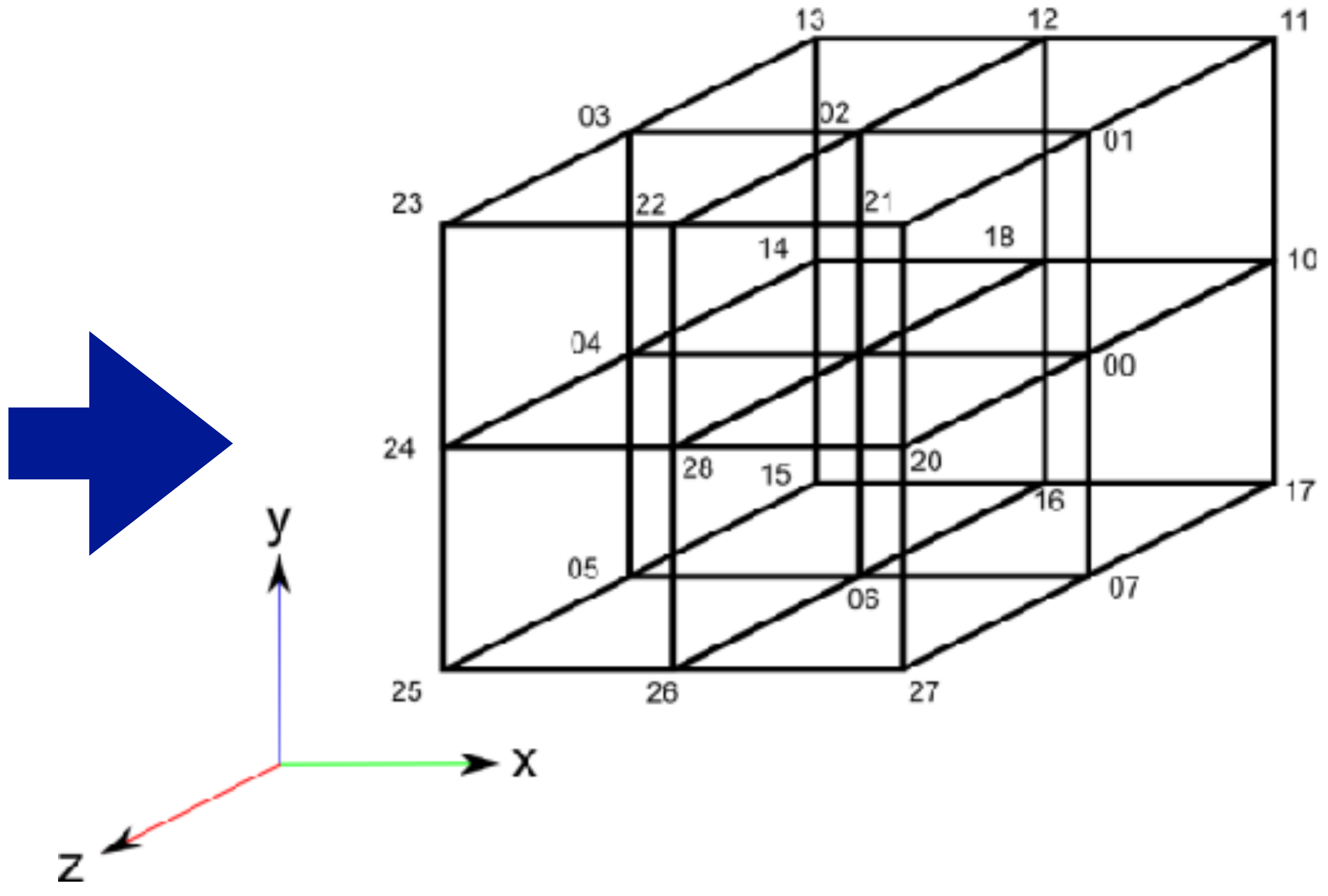
(Atencio, Branch, Sanchez, 2009). Geometric approach for estimating the weight of mango sugar fruit by digital image processing

Visual feature engineering

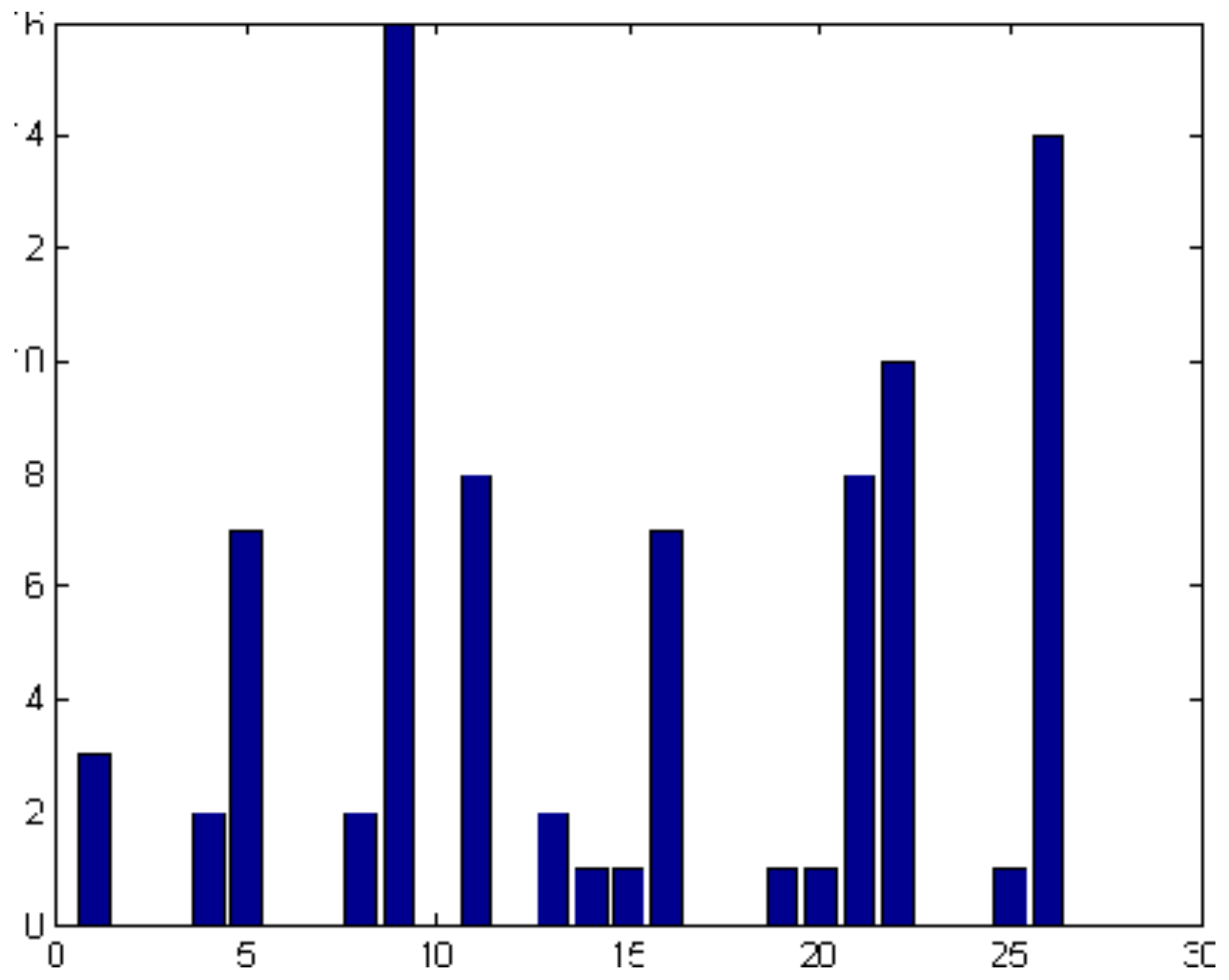
3D Model (hole detection)



3D Bresenham



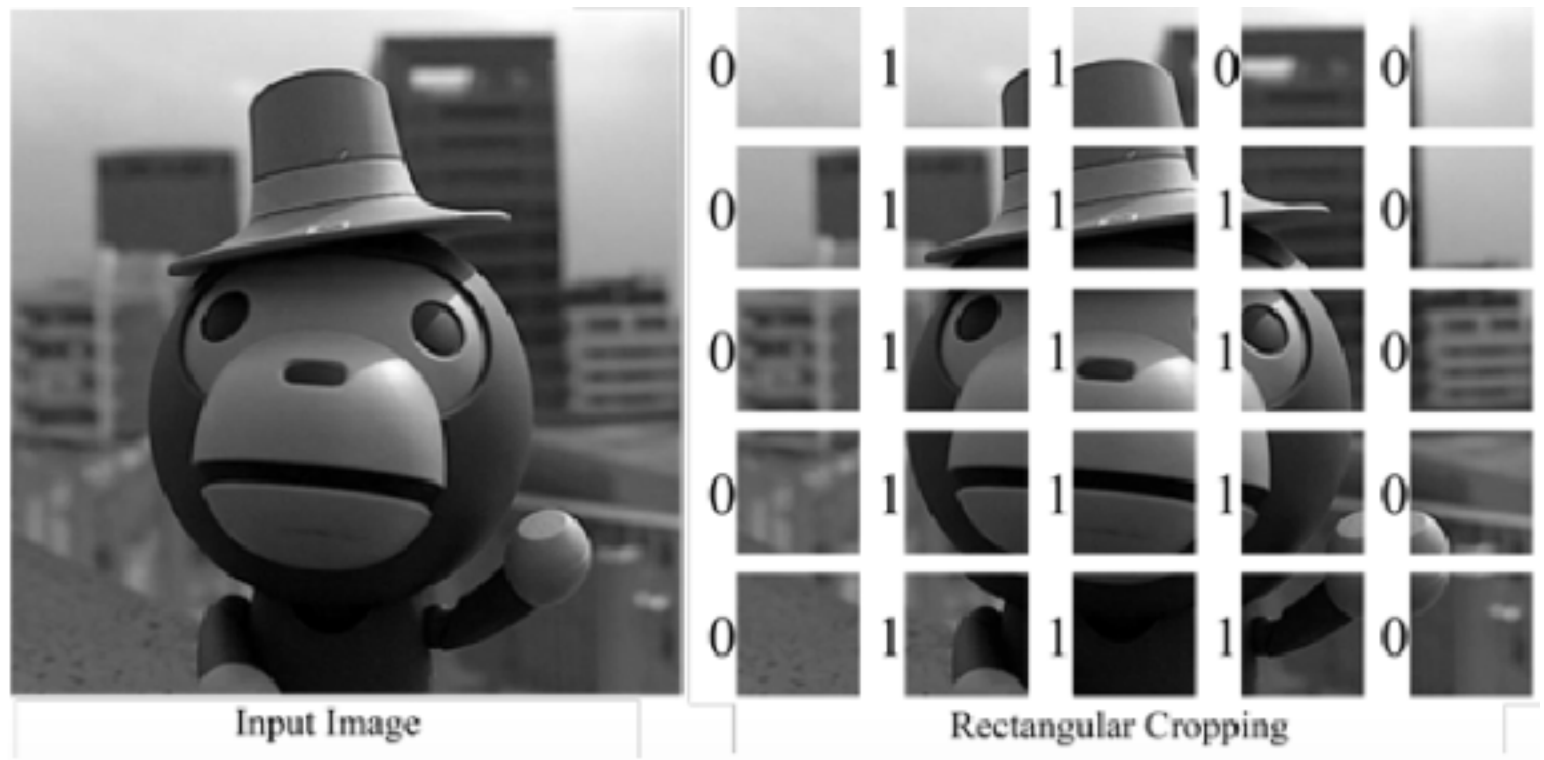
3D Chain-Codes



(Atencio, Branch, Sanchez, 2012). Automatic method for calculation of 3D chain-codes on contours of holes in three dimensional surfaces of free-form objects

Visual feature engineering

Image labelling



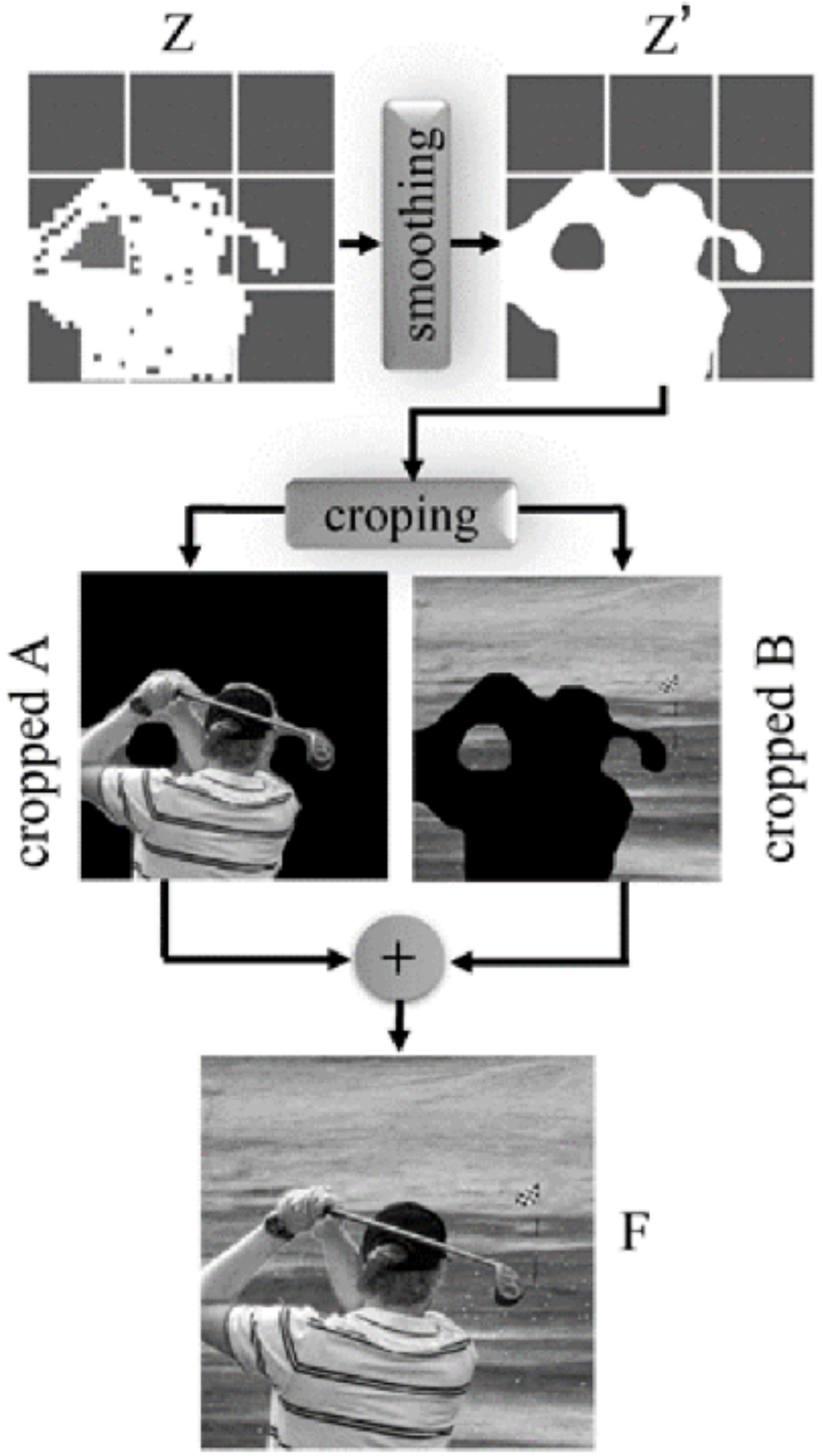
Focus metrics

$$EOL = \sum_{x}\sum_{y}(f_{xx} + f_{yy})^2$$

$$SML = \sum_{i=x-N}^{i=x+N}\sum_{j=x-N}^{j=x+N} \nabla^2{}_{ML}f(i,j) \, for \, \nabla^2{}_{ML}f(i,j) \geq T$$

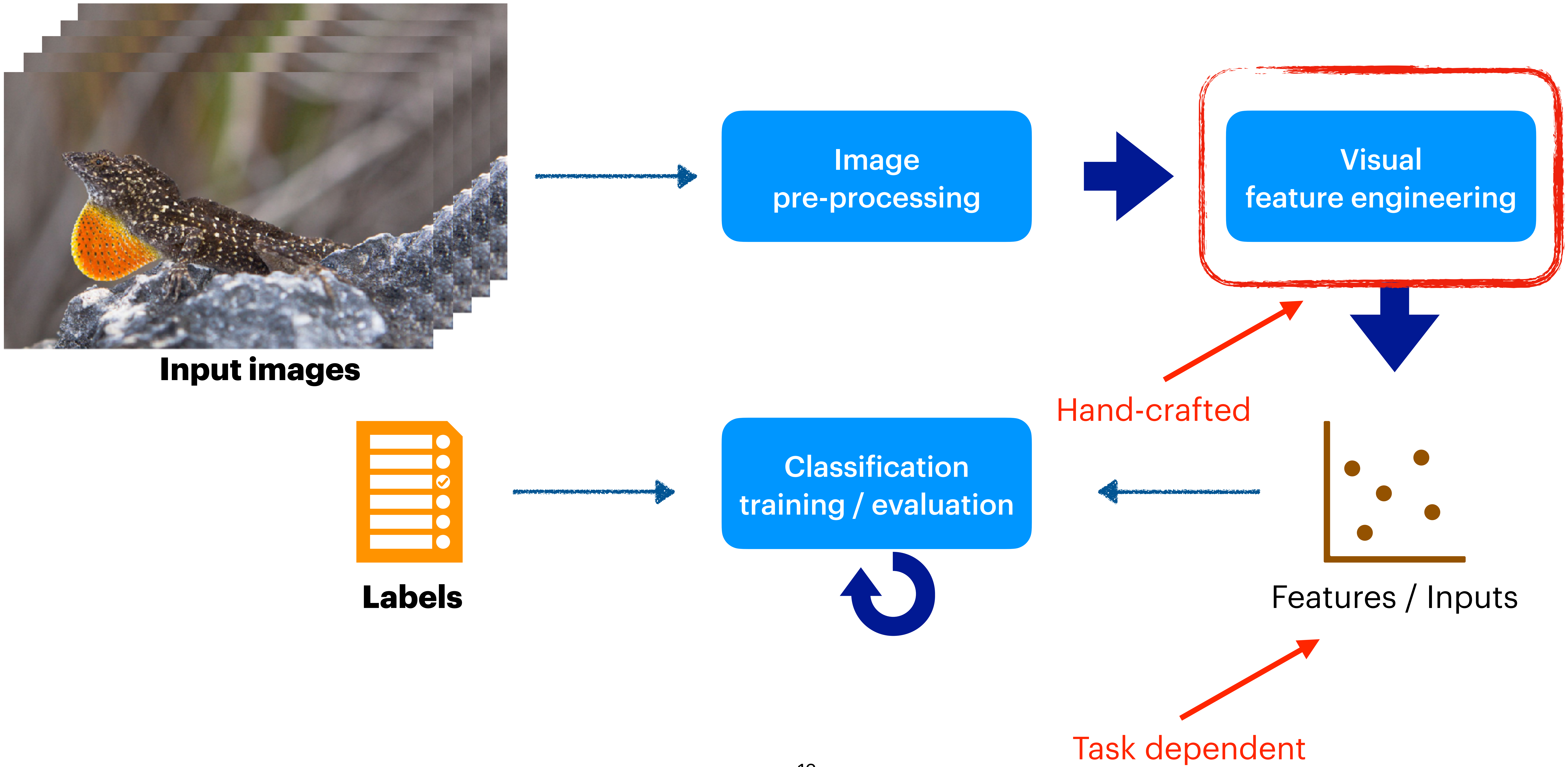
$$EOG = \sum_{x}\sum_{y}(f_x^2 + f_y^2)$$

Multi-focus fusion

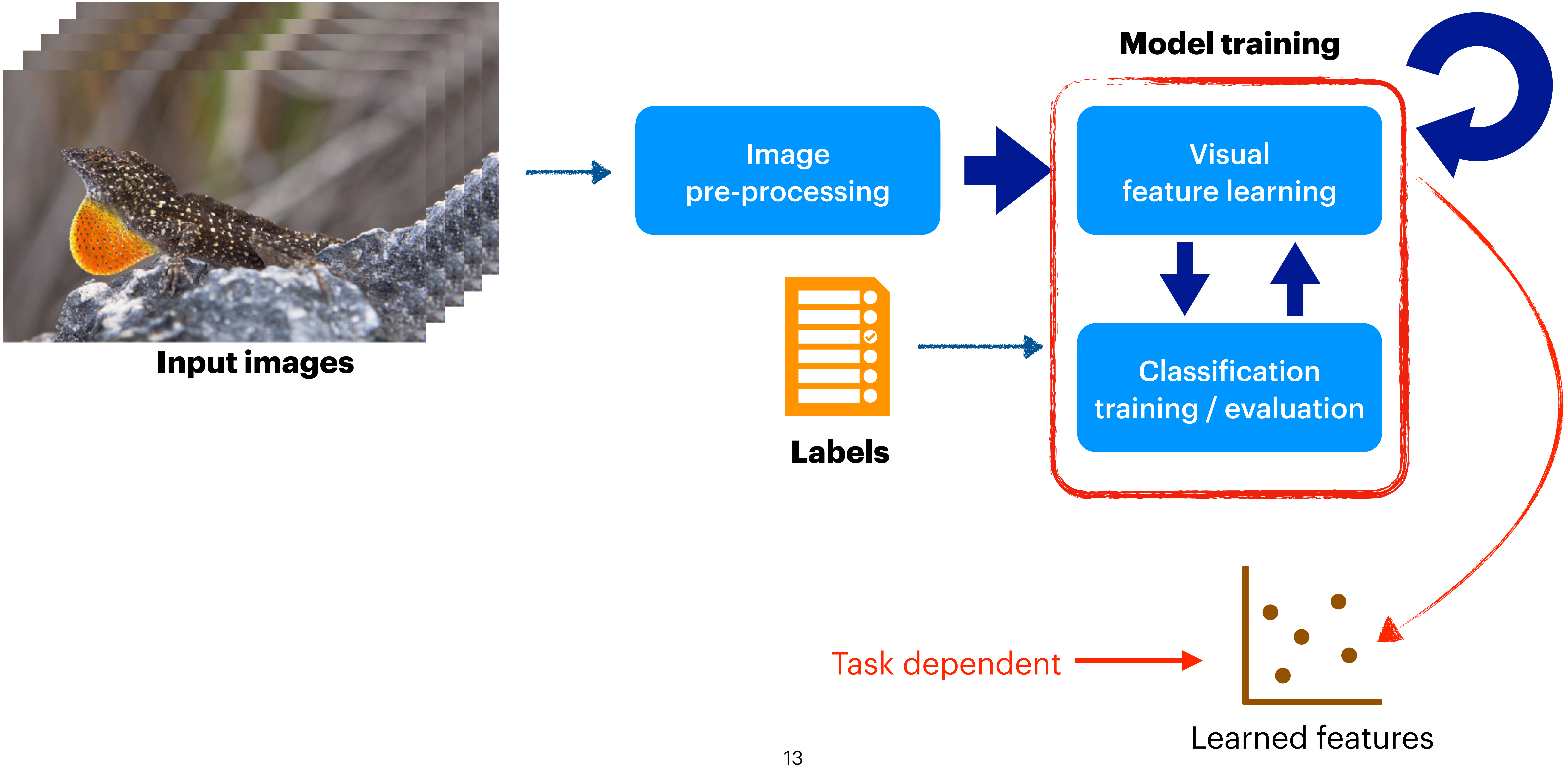


(Atencio, Branch, Sanchez, 2017). Evaluating supervised learning approaches for spatial-domain multi-focus image fusion.

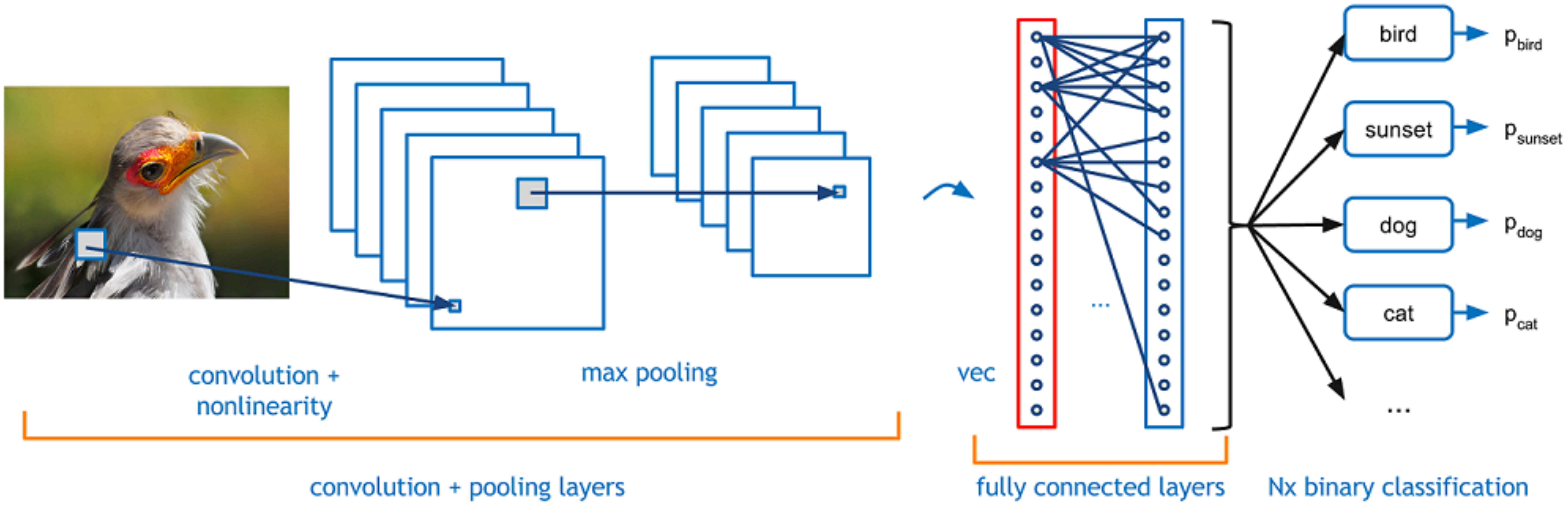
End to end learning



End to end learning

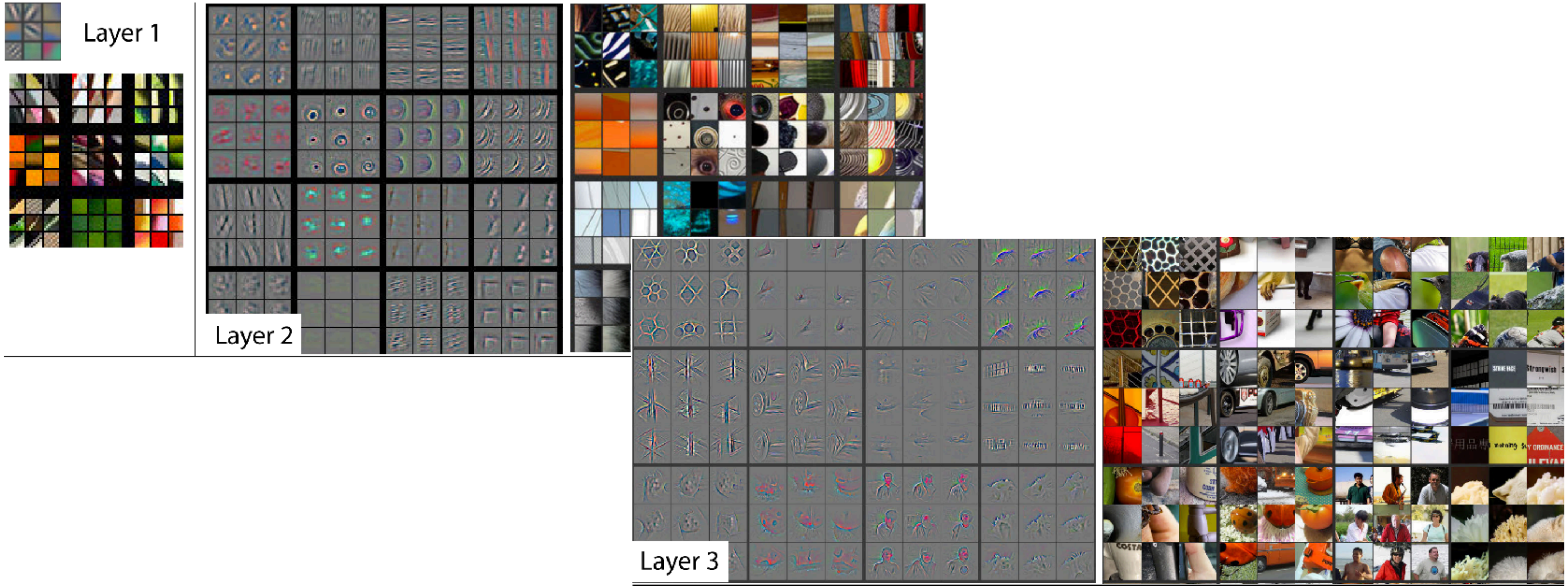


Supervised Representation Learning



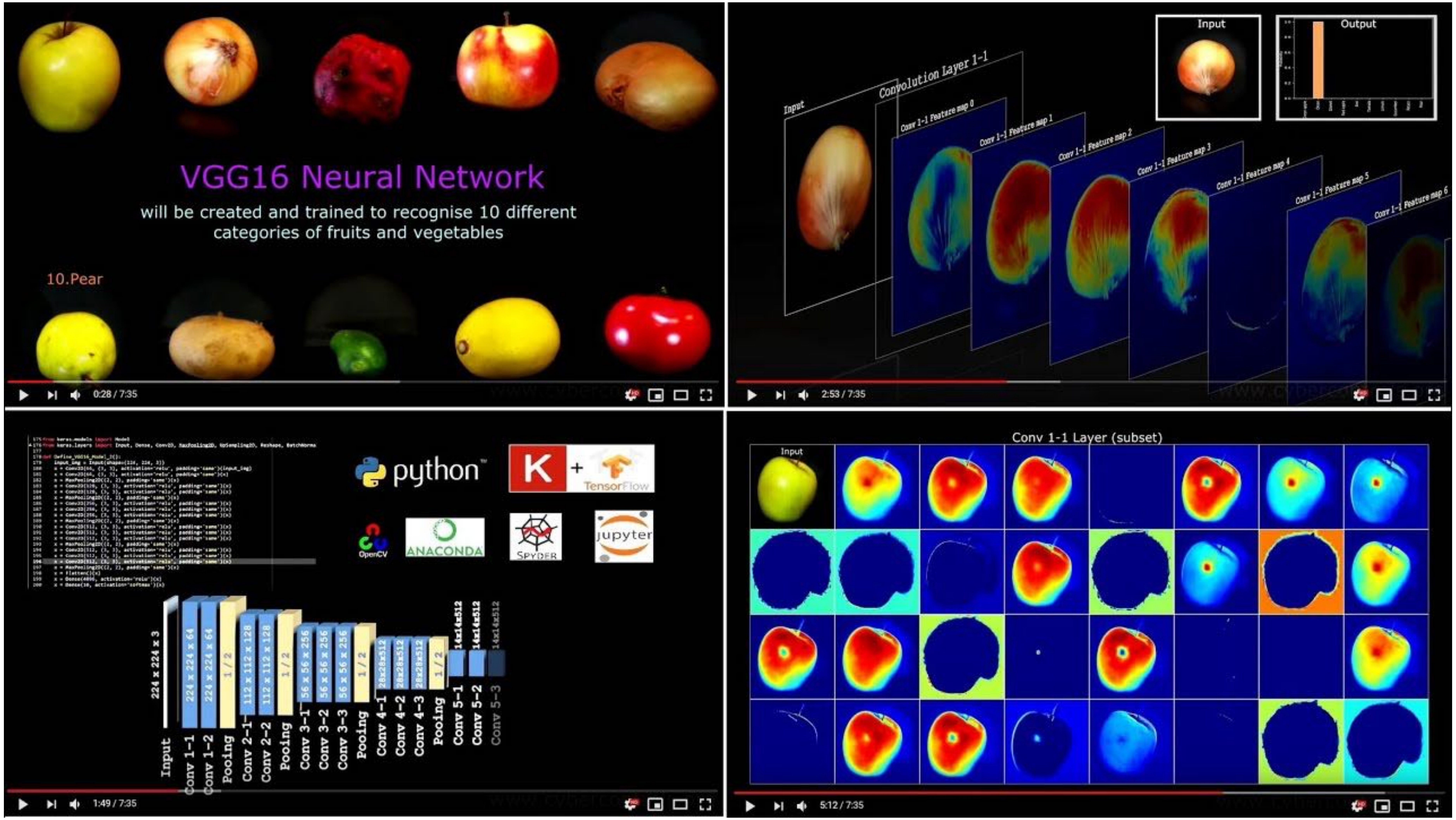
Source: https://adeshpande3.github.io/assets/Cover.png

Supervised Representation Learning

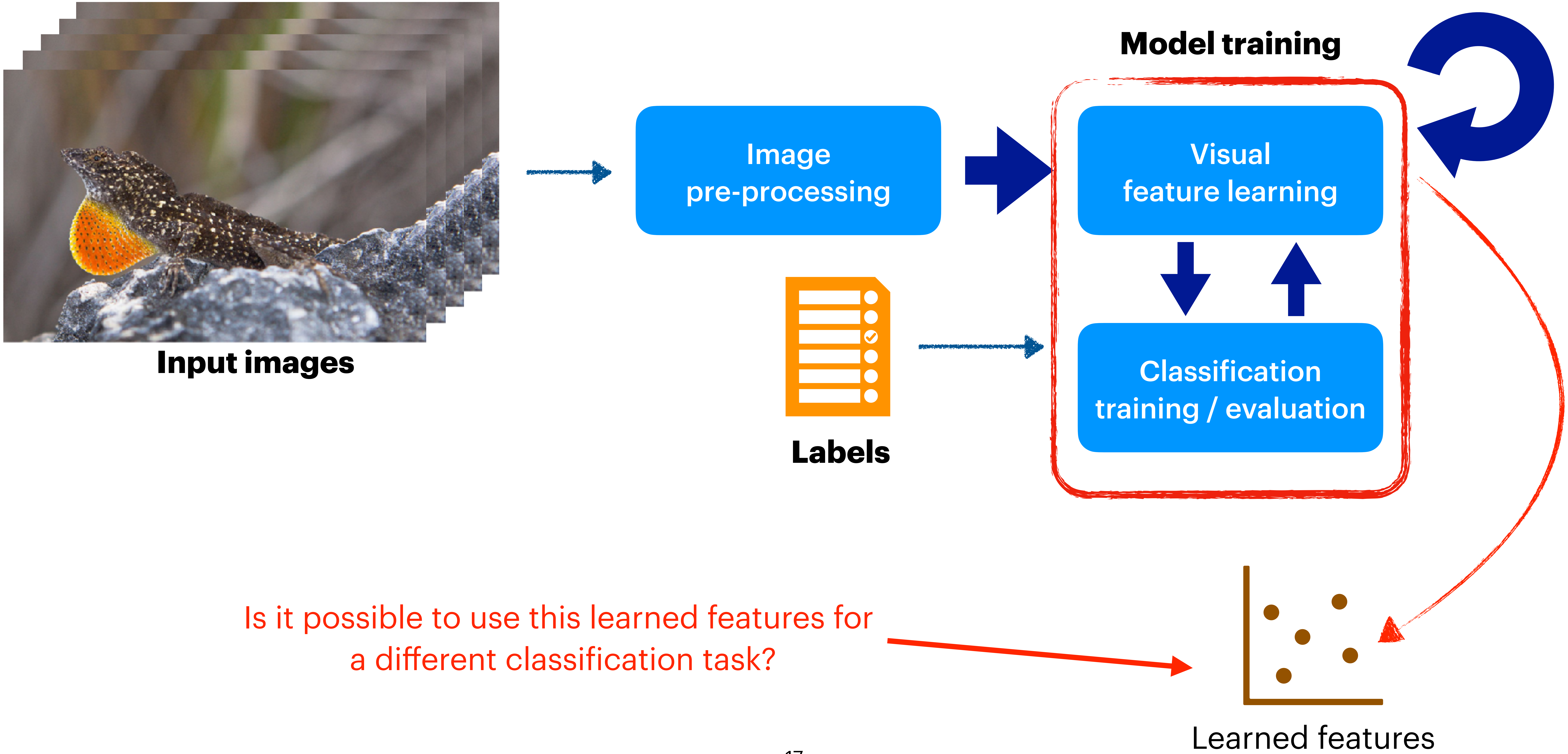


(Zeiler & Fergus, 2013). Visualizing and Understanding Convolutional Networks.

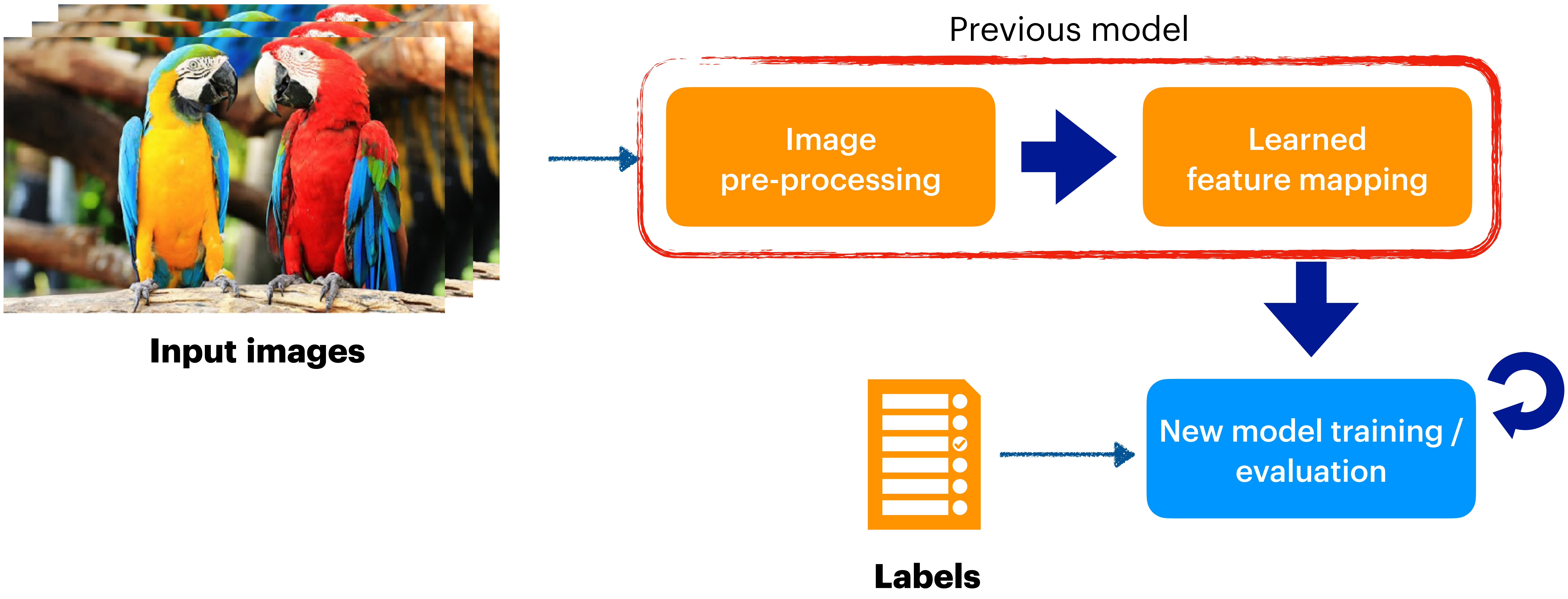
Supervised Representation Learning



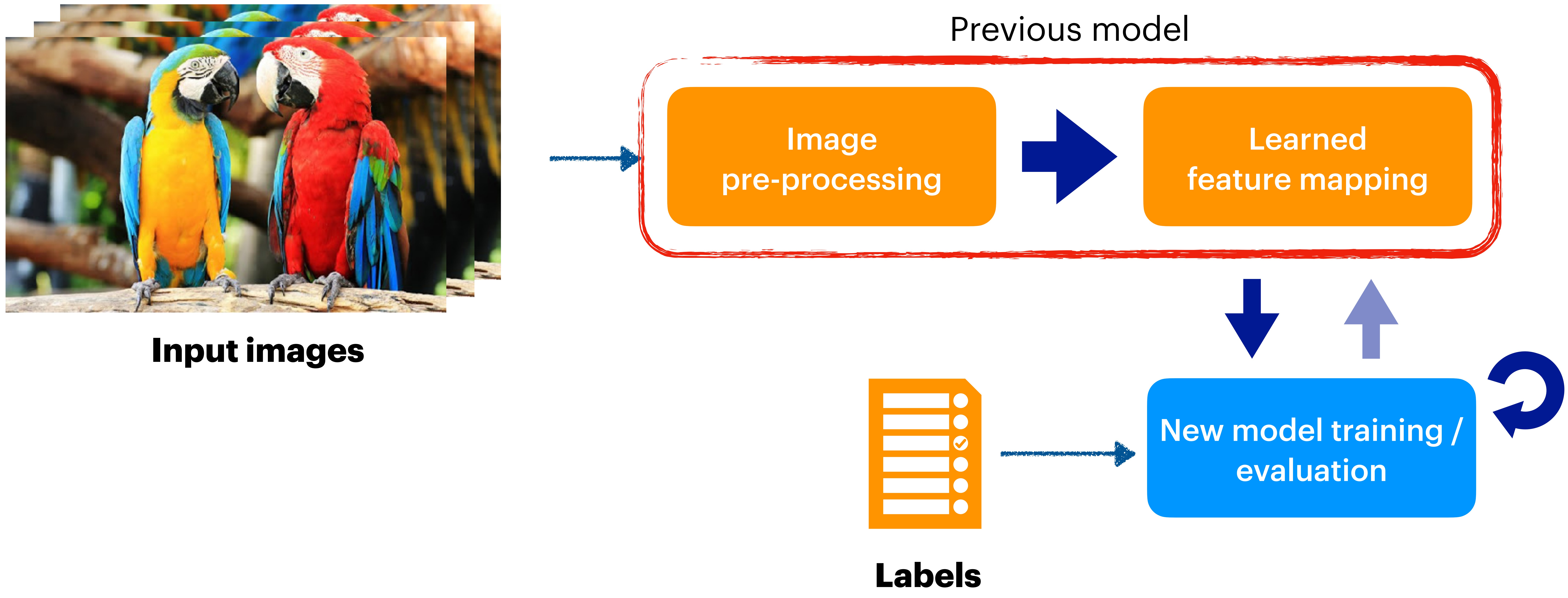
Transfer Learning



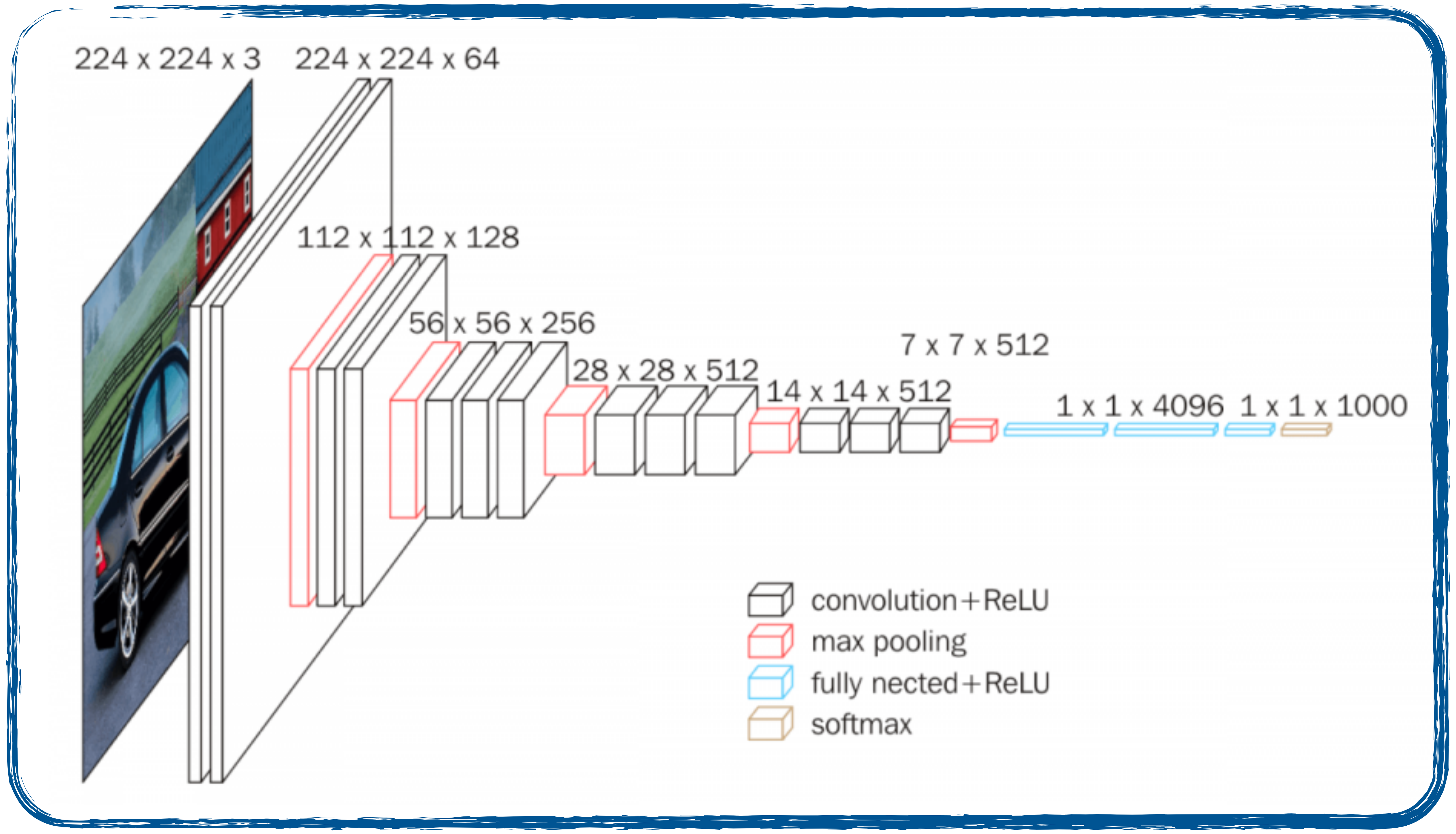
Transfer Learning



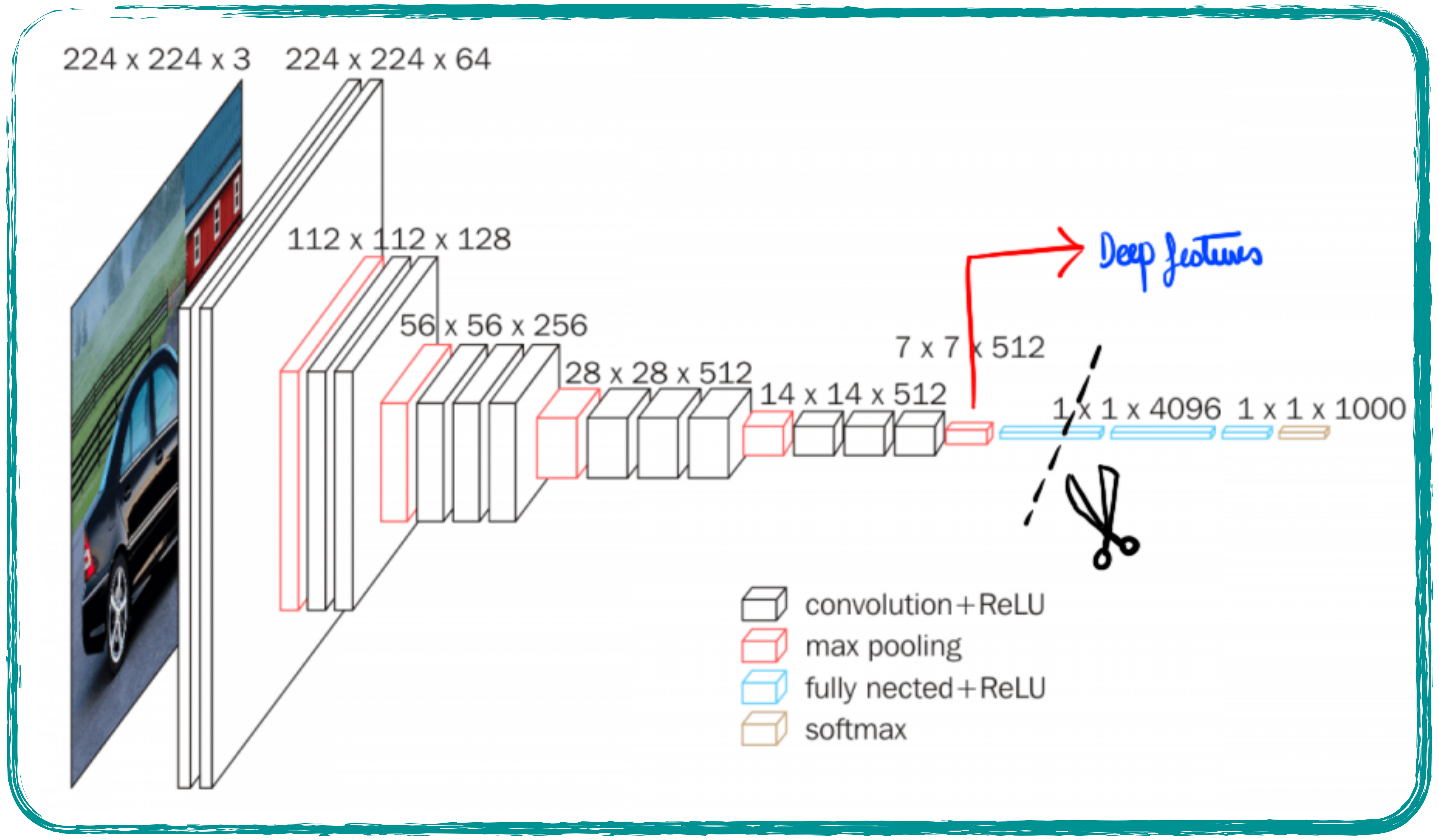
Transfer Learning + fine tuning



Transfer Learning + fine tuning

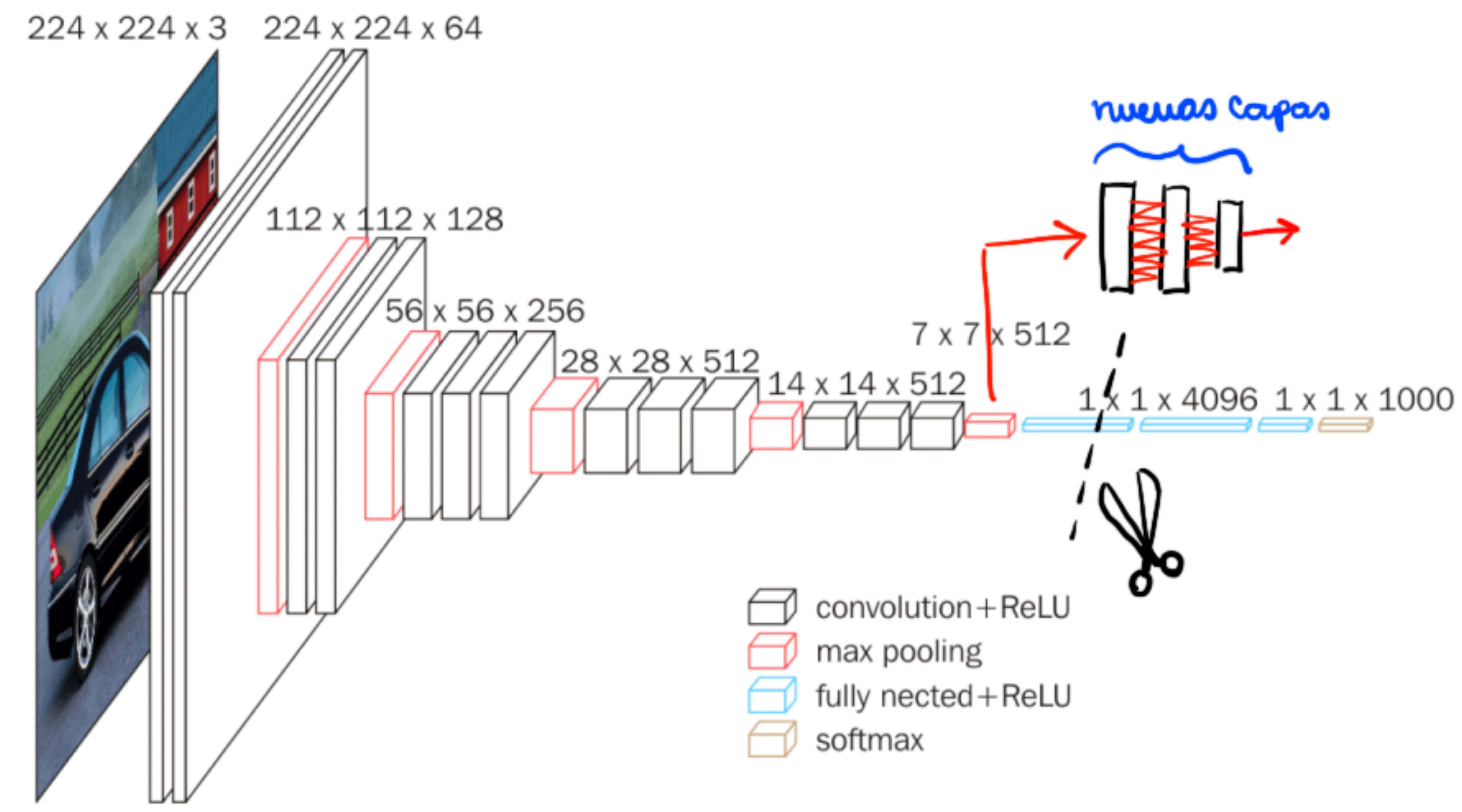


Model A



Learned features from Model A

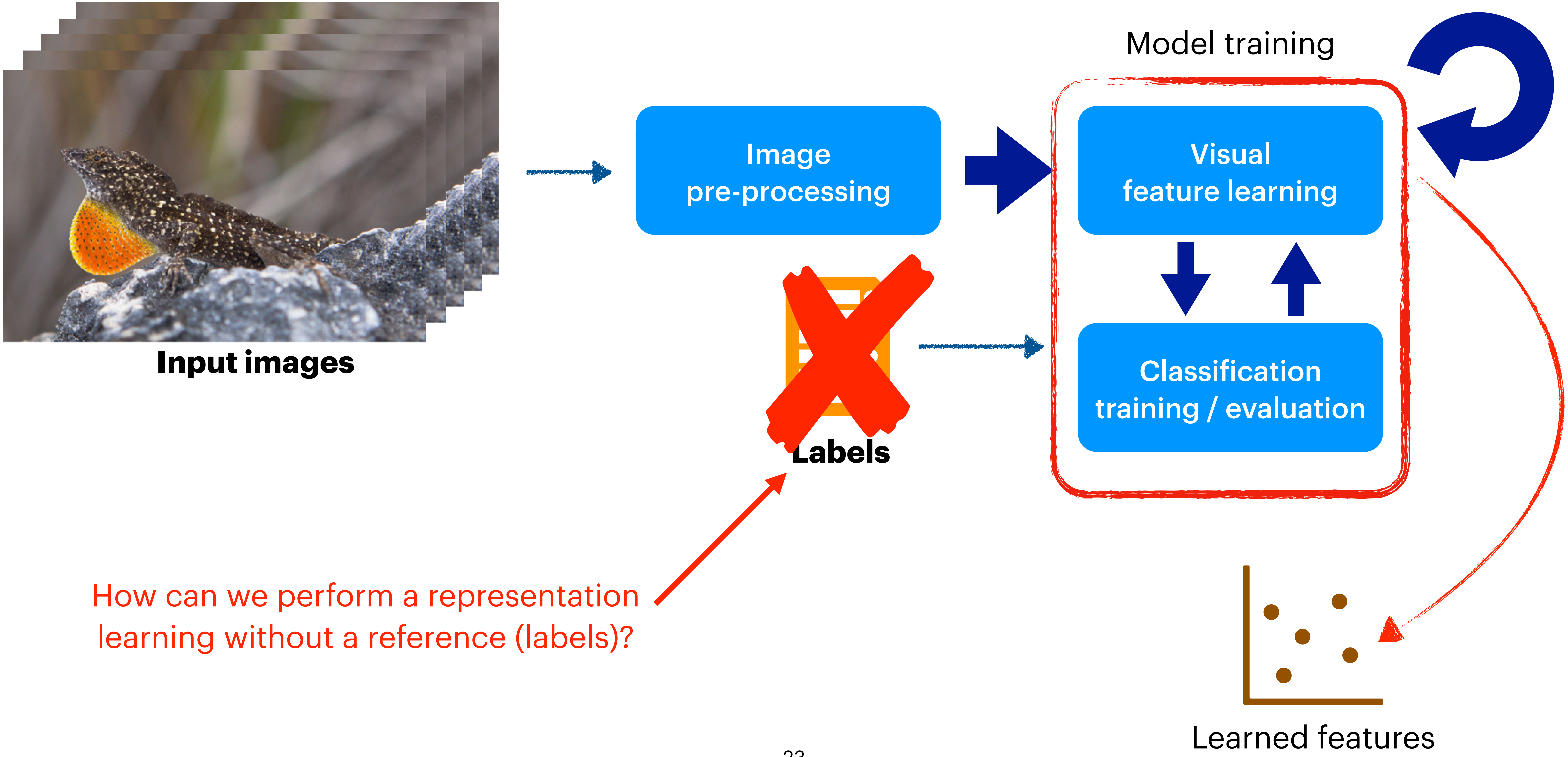
Transfer Learning + fine tuning



Features from ModelA + Model B

2. The unsupervised representation learning problem

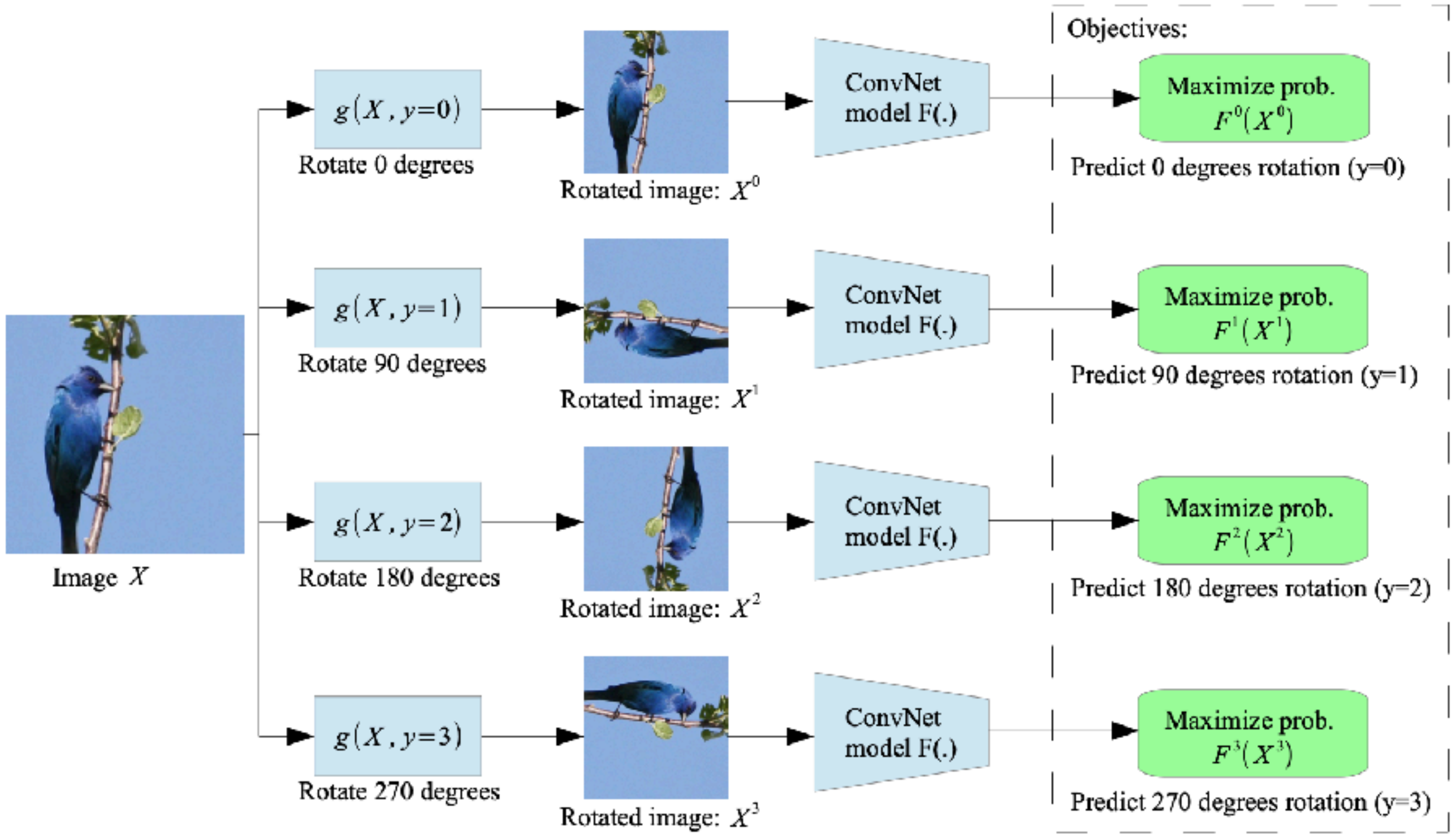
Representation learning: unsupervised case



How can we perform a representation learning without a reference (labels)?

Self-supervised learning

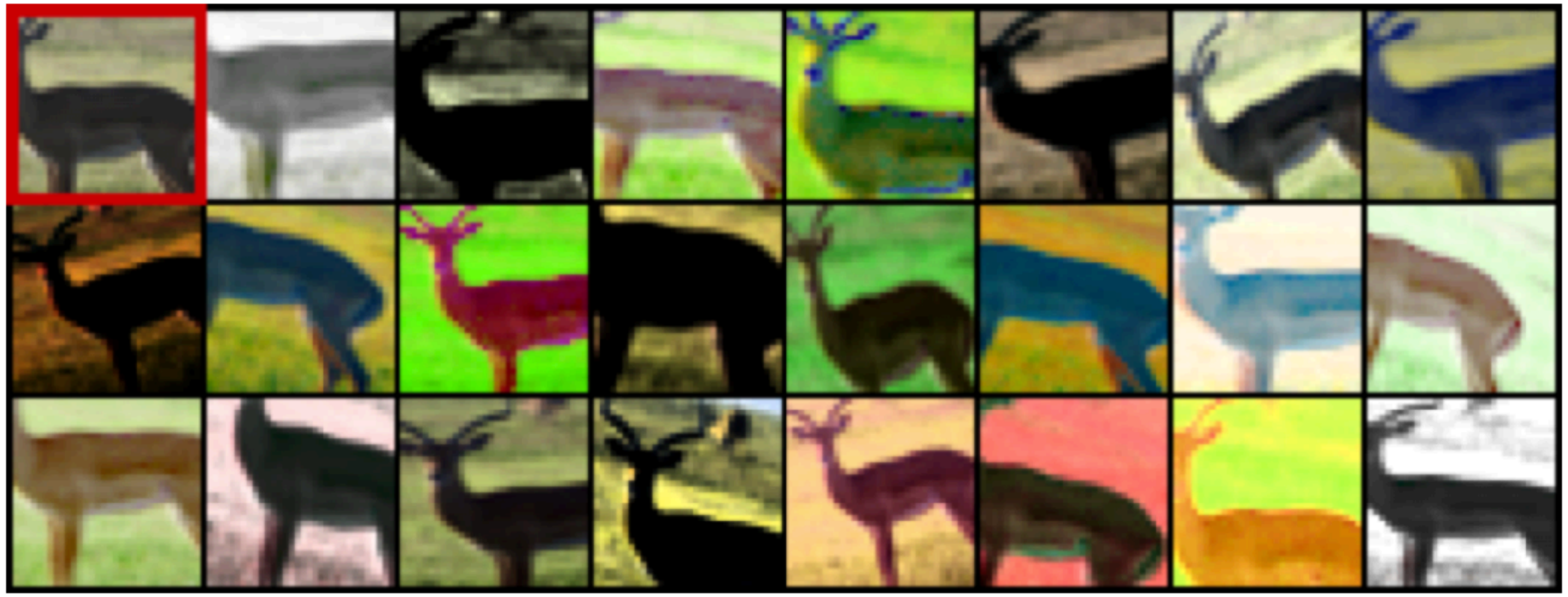
- Recent Machine Learning Approach.
- Takes concepts from Natural Language Processing.
- Introduces the concept of **pretext task**.
- Does not requires labelling, just input data.
- **Main idea**: Use a pretext task which depends only of the input data, in order to learn a representation of the input.



(Gidaris , et.al. 2018). UNSUPERVISED REPRESENTATION LEARNING BY PREDICTING IMAGE ROTATIONS.

Self-supervised learning

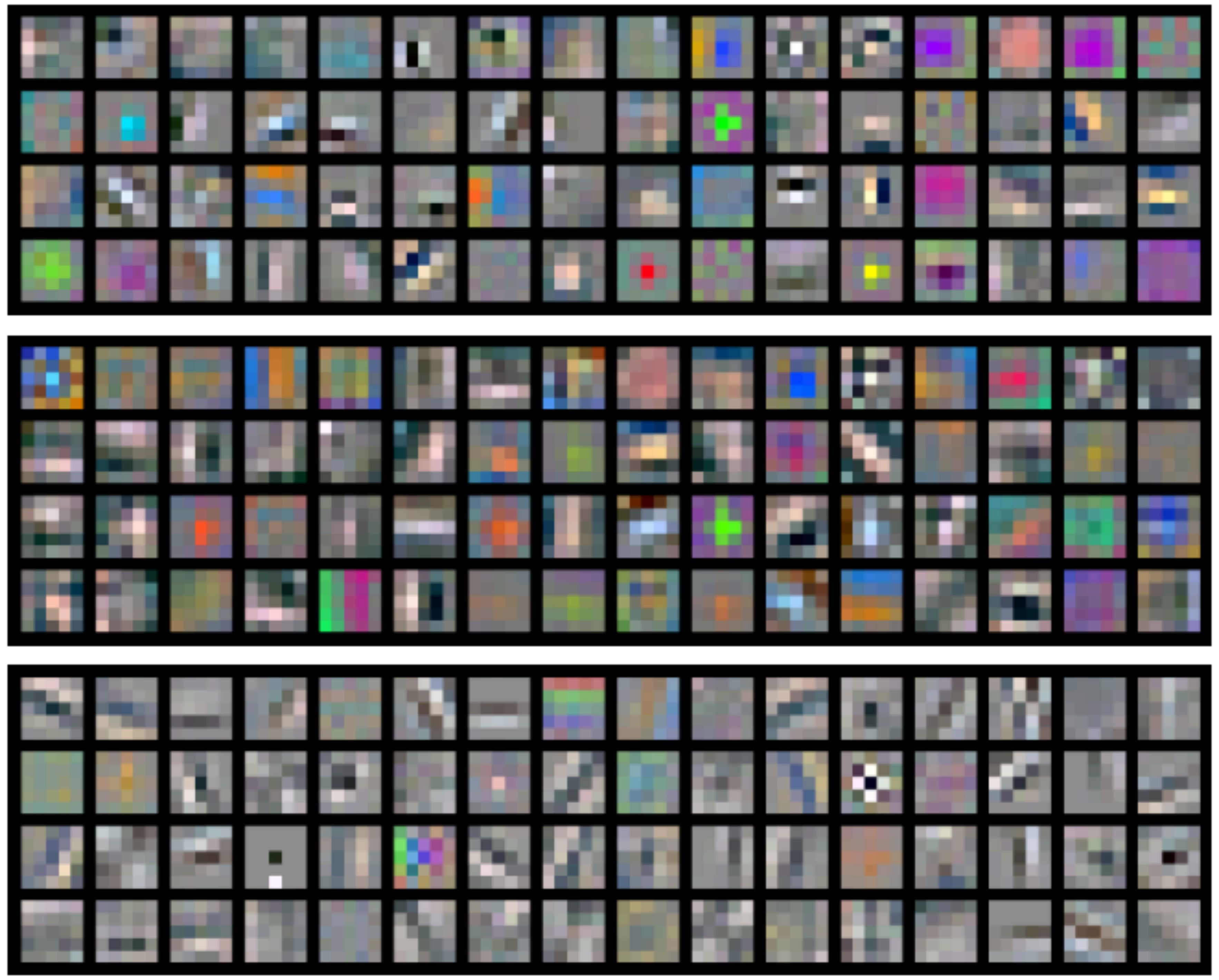
- **Pretext task**: identify an image independent of a set pf transformation applied to it.
- Each image is a class.
- For each image a random transformation is applied.
- Transformations can be: translation, scaling, rotation, contrast modification, color modification.



(Dosovitskiy et.al. 2015). Discriminative Unsupervised Feature Learning with Exemplar Convolutional Neural Networks.

Self-supervised learning

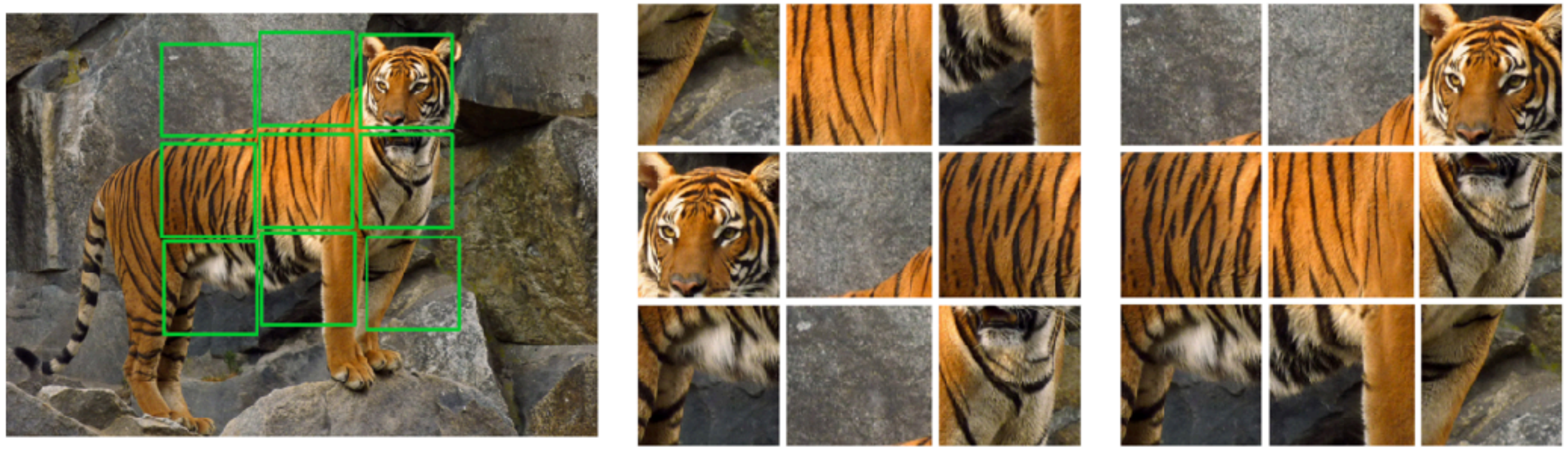
- **Pretext task**: identify an image independent of a set pf transformation applied to it.
- Each image is a class.
- For each image a random transformation is applied.
- Transformations can be: translation, scaling, rotation, contrast modification, color modification.



(Dosovitskiy et.al. 2015). Discriminative Unsupervised Feature Learning with Exemplar Convolutional Neural Networks.

Self-supervised learning

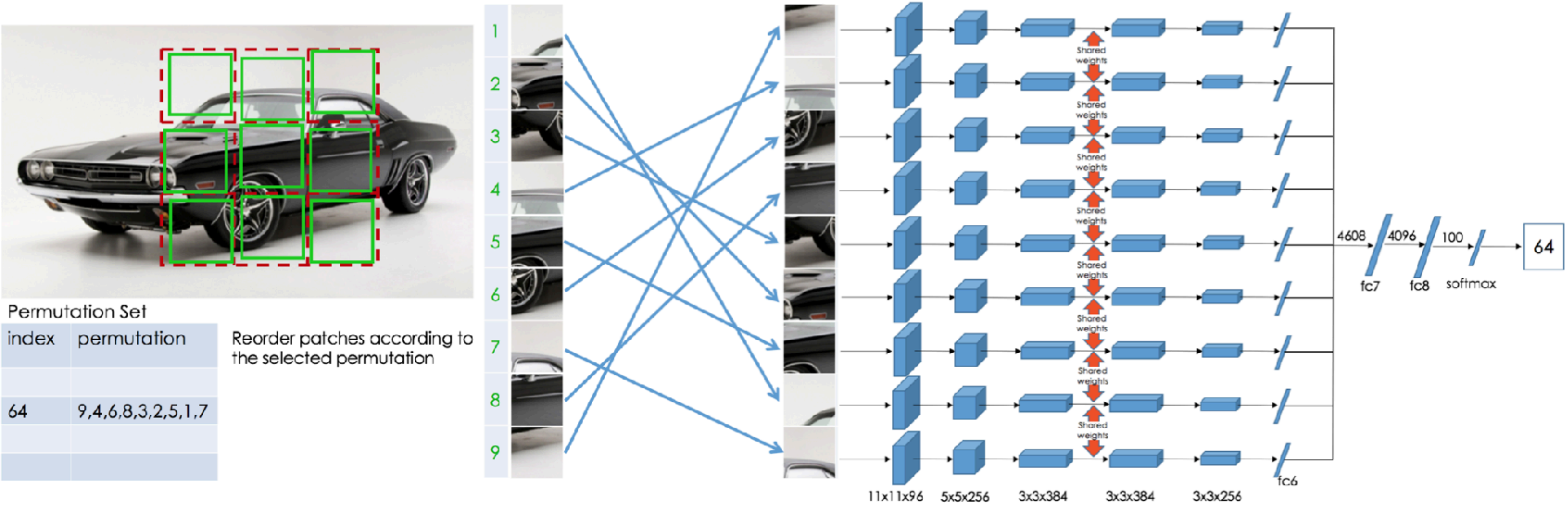
- **Pretext task**: To predict the index of the chosen permutation from a set of 64 pre-defined permutations.
- 9-puzzle of patches 64x64 pixels.
- The network 9 images randomly positioned and predicts how they are organized.
- 64 classes classification problem.



(a) (b) (c)

(Noroozi & Favaro. 2017). Unsupervised Learning of Visual Representations by Solving Jigsaw Puzzles.

Self-supervised learning



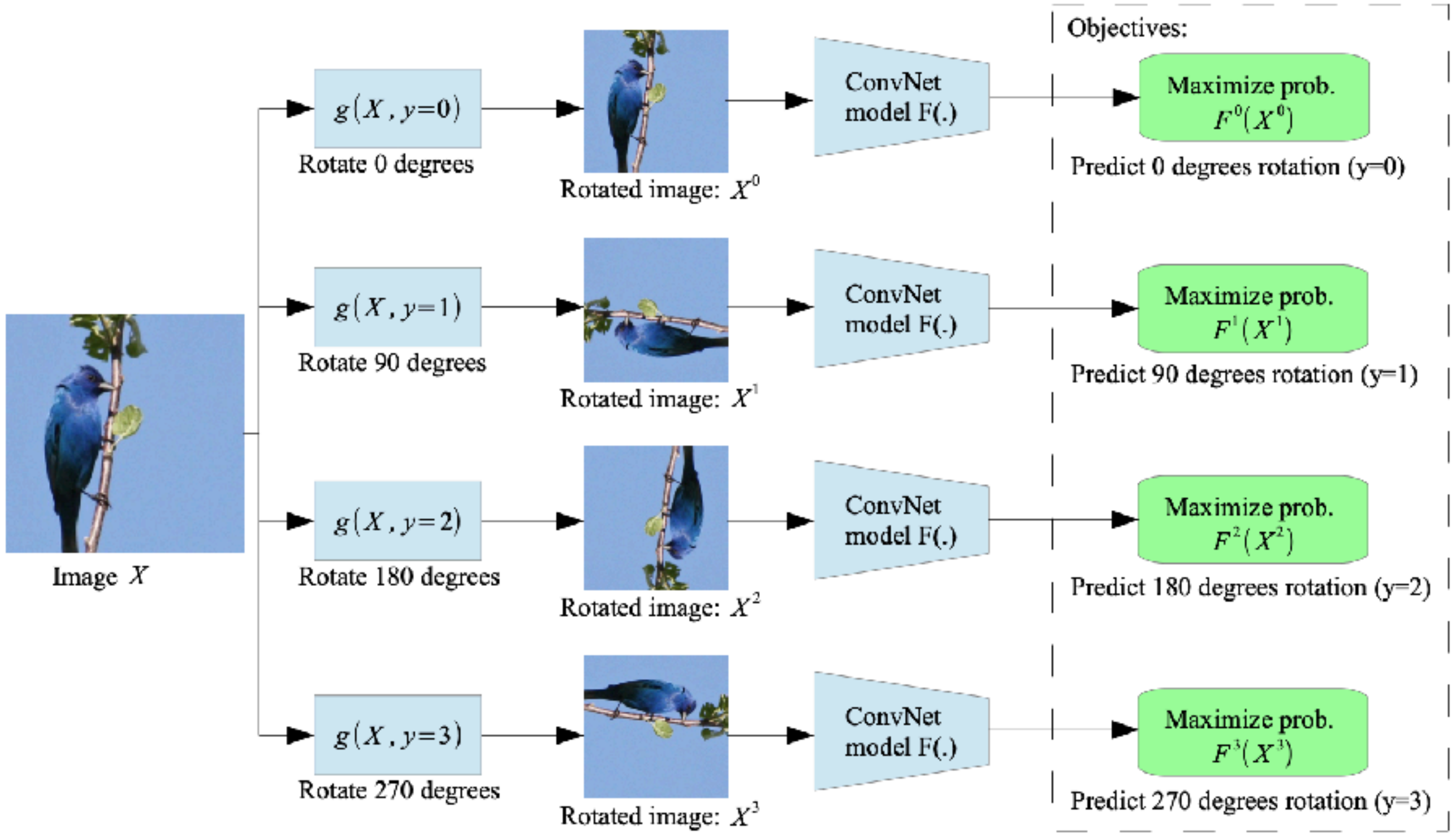
(Noroozi & Favaro. 2017). Unsupervised Learning of Visual Representations by Solving Jigsaw Puzzles.

Self-supervised learning

- **Pretext task**: To predict the rotation applied to an input image.
- Each image can be rotated by 0, 90, 180 or 270 degrees.
- A feed-forward convolutional network is used.
- 4 classes classification problem.

(Gidaris , et.al. 2018). UNSUPERVISED REPRESENTATION LEARNING BY PREDICTING IMAGE ROTATIONS.

Data generation

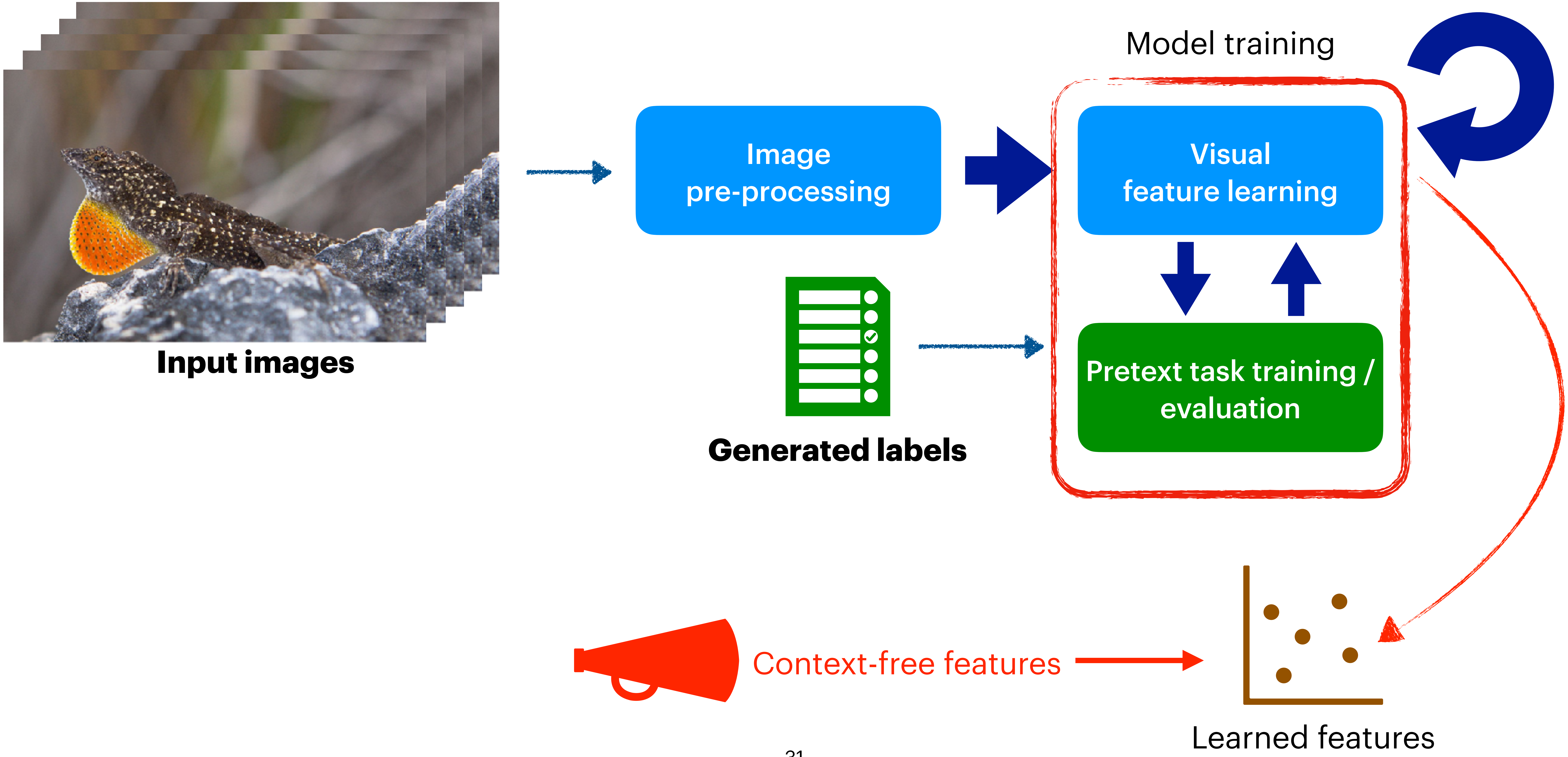


Self-supervised learning

(Gidaris , et.al. 2018). UNSUPERVISED REPRESENTATION LEARNING BY PREDICTING IMAGE ROTATIONS.

(a) **Attention maps of supervised model** (b) **Attention maps of our self-supervised model**

Self-representation learning



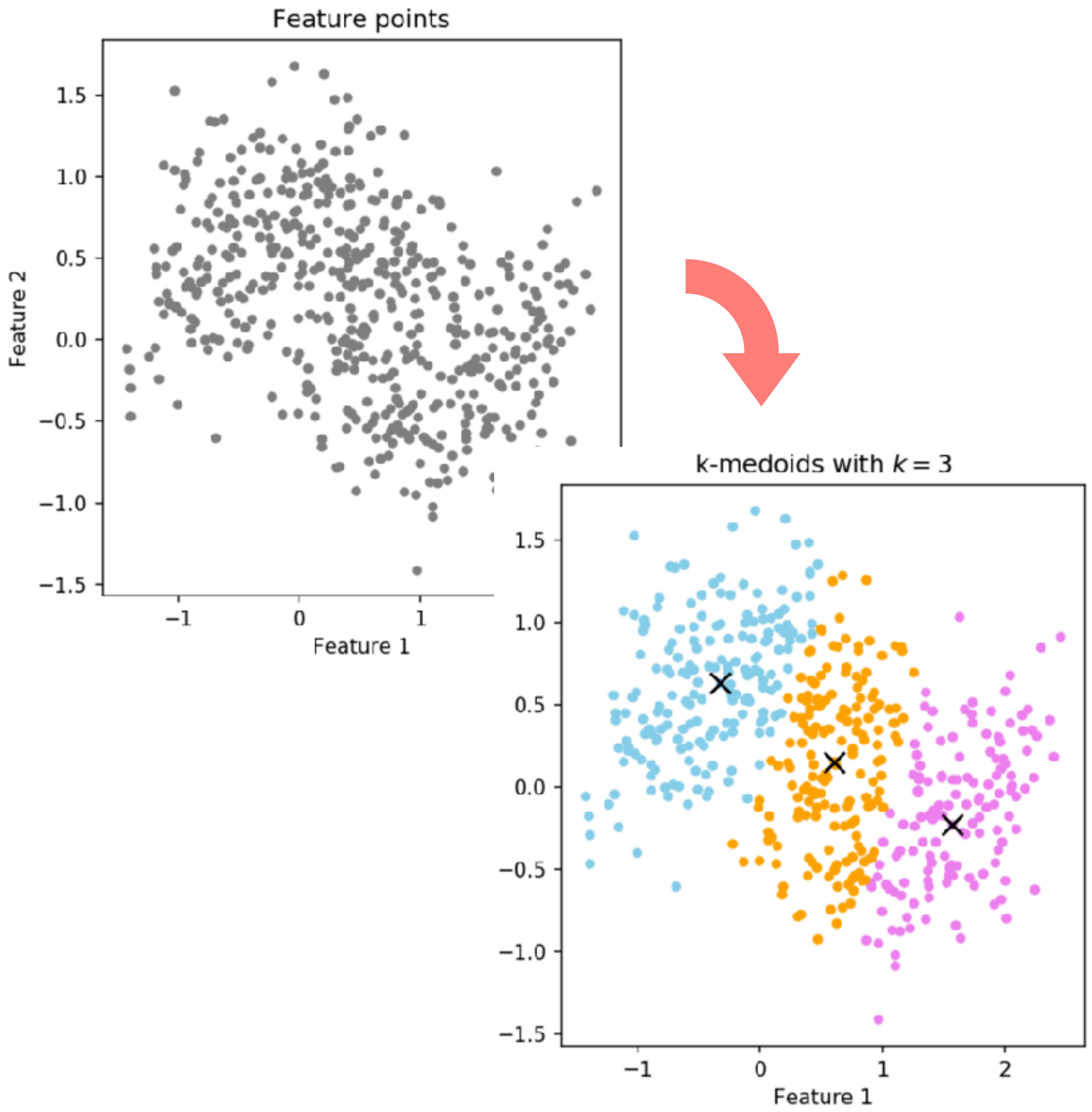
3. Clustering for unlabelled classification

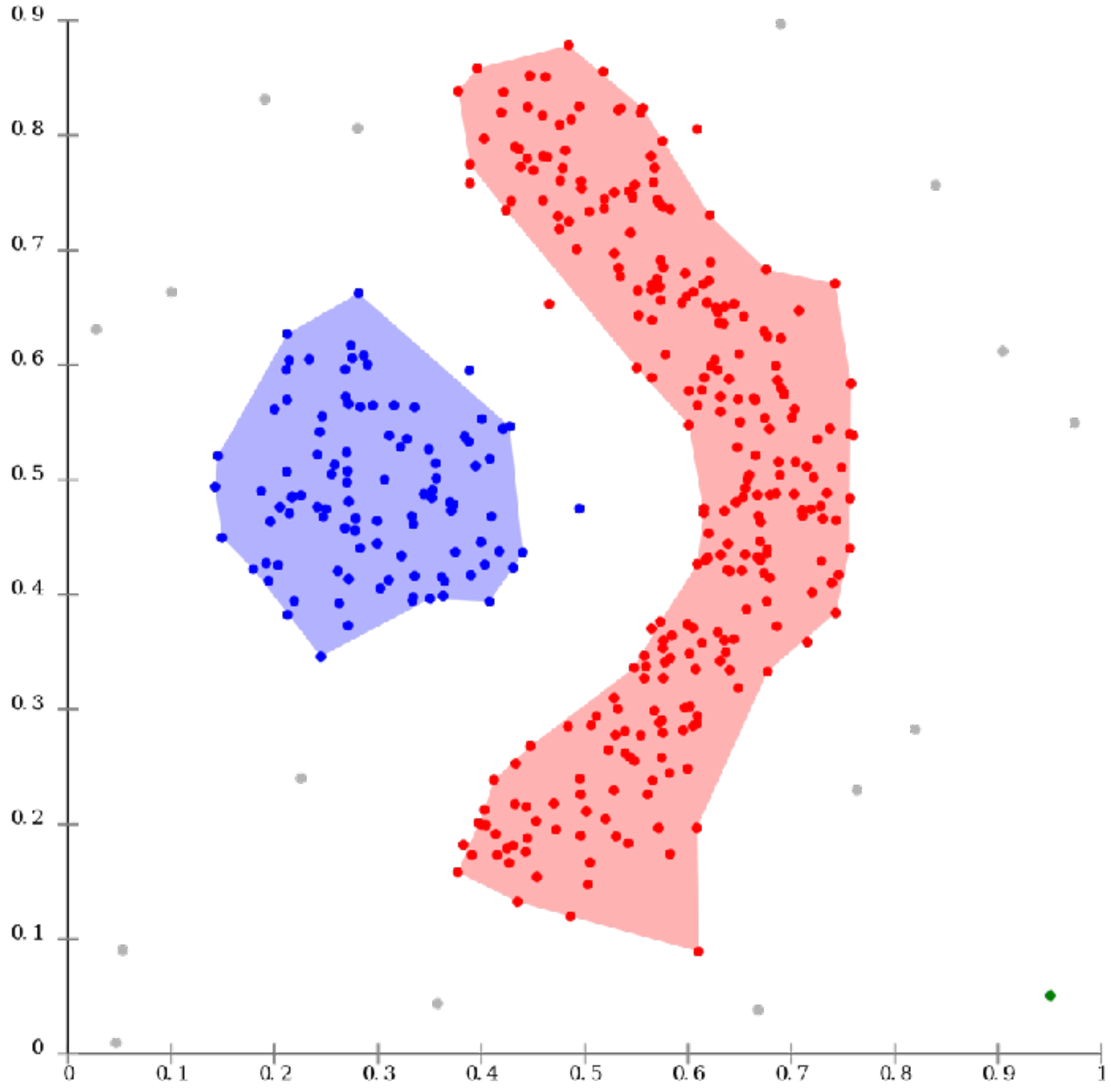
Clustering for label identification

- Once we have a feature space or representation for input data, we can use clustering in order to identify the samples belonging to a set.
- Feature space can be hand-crafted, transferred or self-learned.
- Using a cluster or distance approach we can assign a label to a new sample.
- Parametric or non-parametric techniques can be employed.

(Otani et. al. 2016) Video Summarization using Deep Semantic Features.

Classical clustering

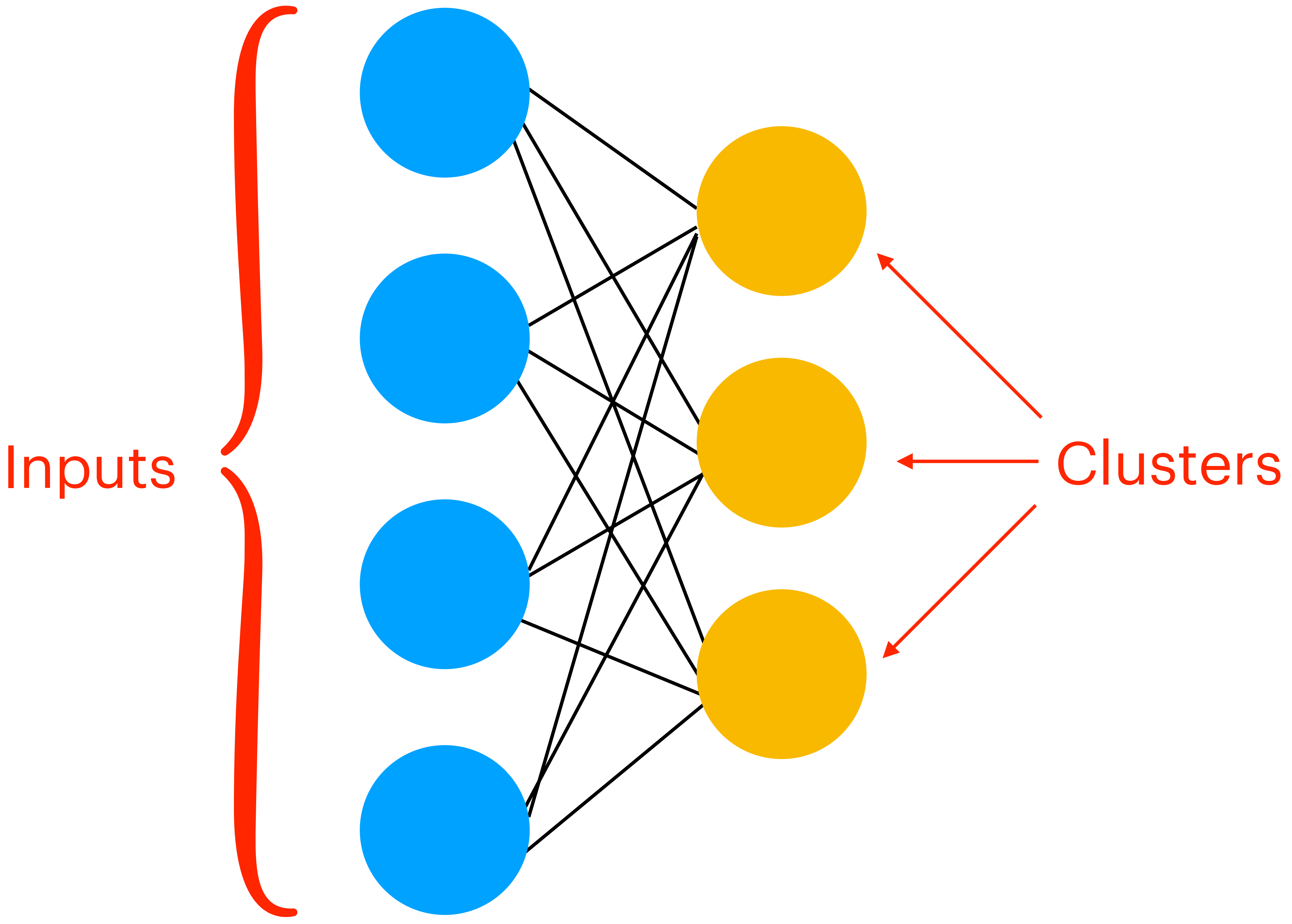




K-Means as a MLP

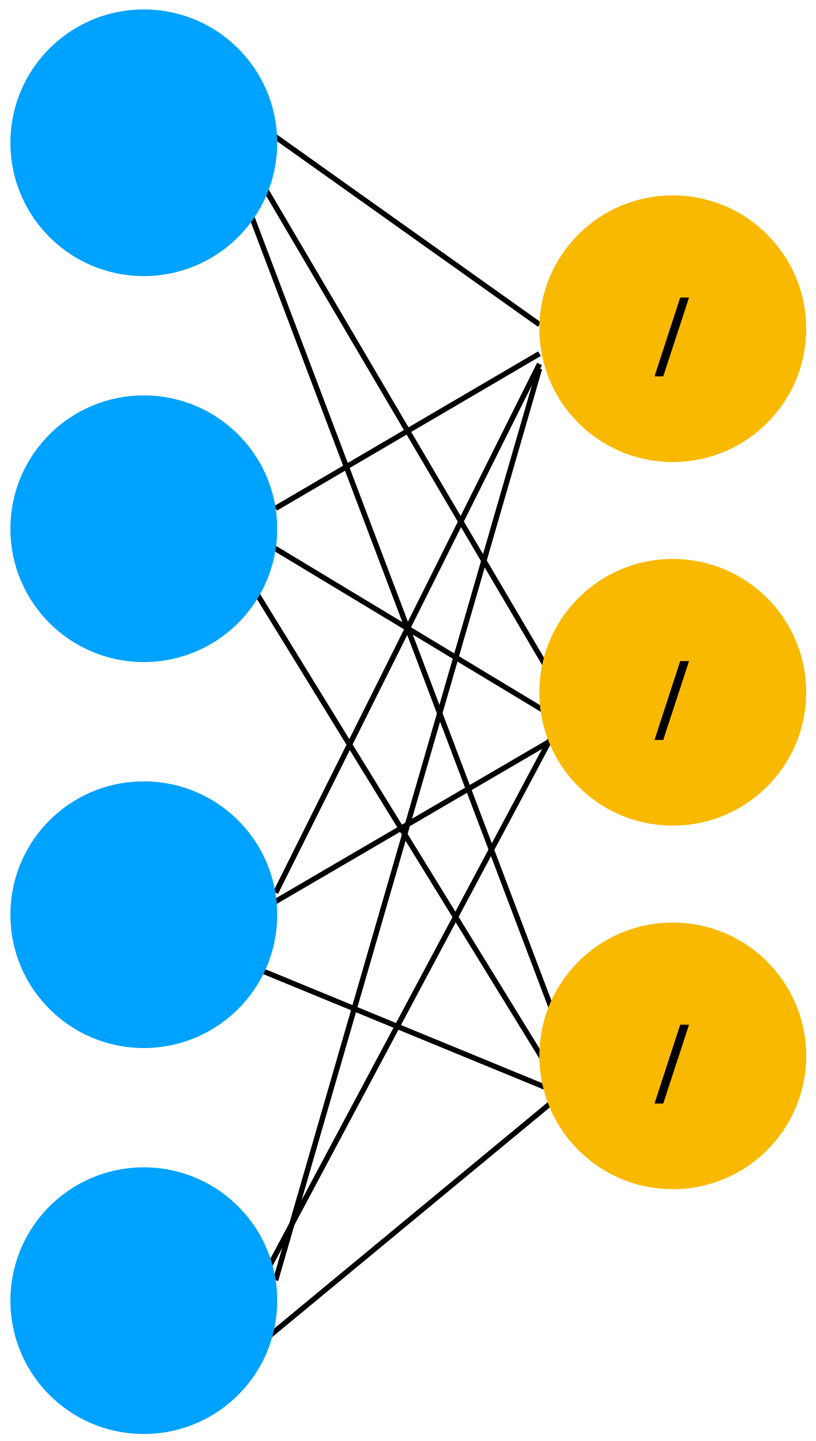
- It is possible to pose a clustering problem as a Neural Network optimization.
- This will allow the system to be complete using a single neural-based architecture.
- Clusters can be associated with the network parameters.
- Loss function can be associated with the intra-cluster cohesion (SSE).

$$SSE(C_i) = \sum_{x \in C_i} d(C_i, x)^2 = \frac{1}{2m} \sum_{x \in C_i} \sum_{y in C_i} d(x, y)^2$$



K-Means as a MLP

Given an input vector $x_i = \{x_i^{(1)}, x_i^{(2)}, \ldots, x_i^{(n_x)}\}$ with n_x features, initialize a matrix of weights $W \in \mathbb{R}^{(n_x,k)}$ where each $W_{(:,i)}$ is the features of cluster i for k clusters.



Output can be computed as:

$a_i = W^T \cdot x_i$

Algorithm

Repeat for *m* epochs:

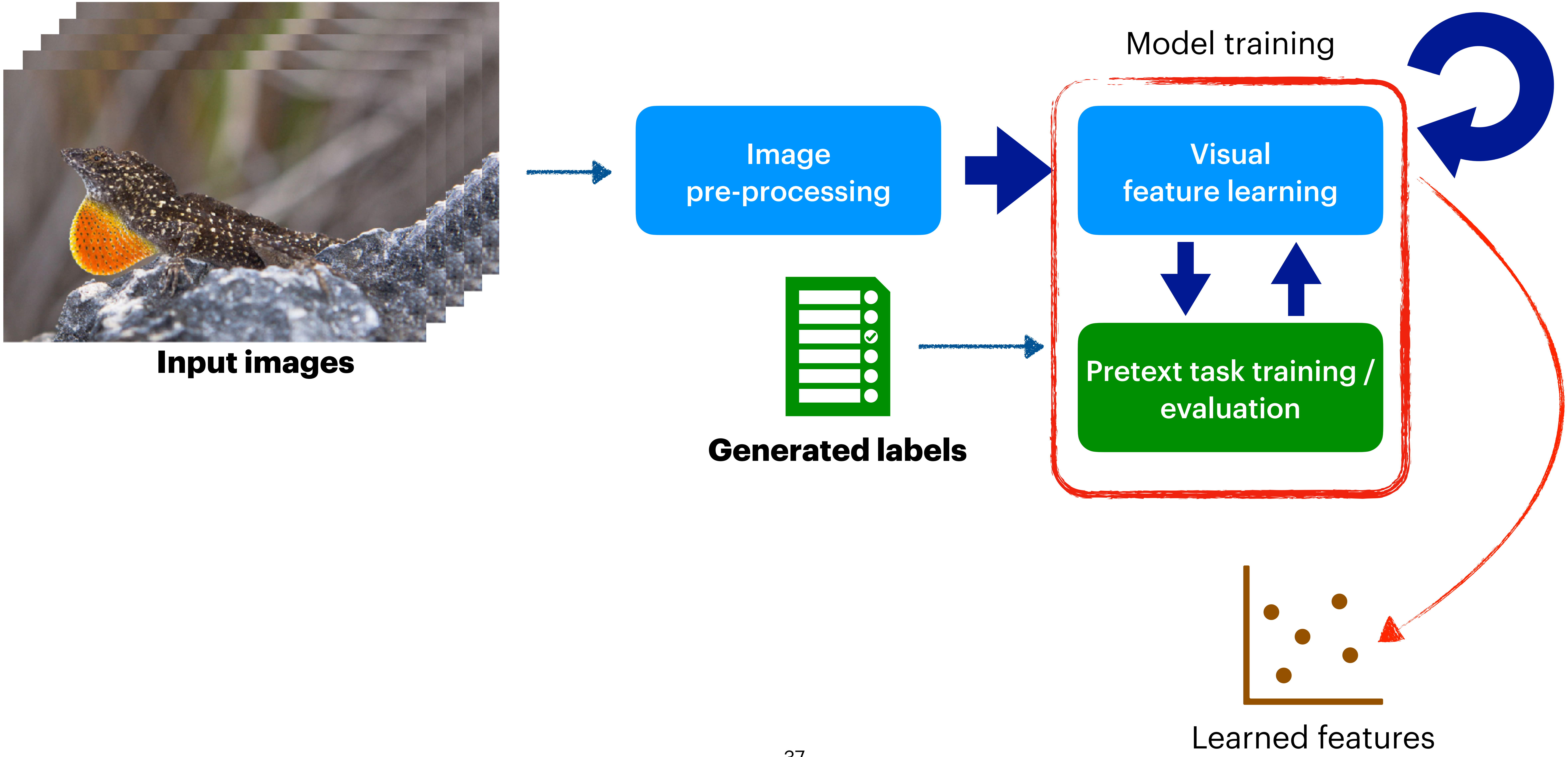
for each data_point x_i:

Compute activations a_i

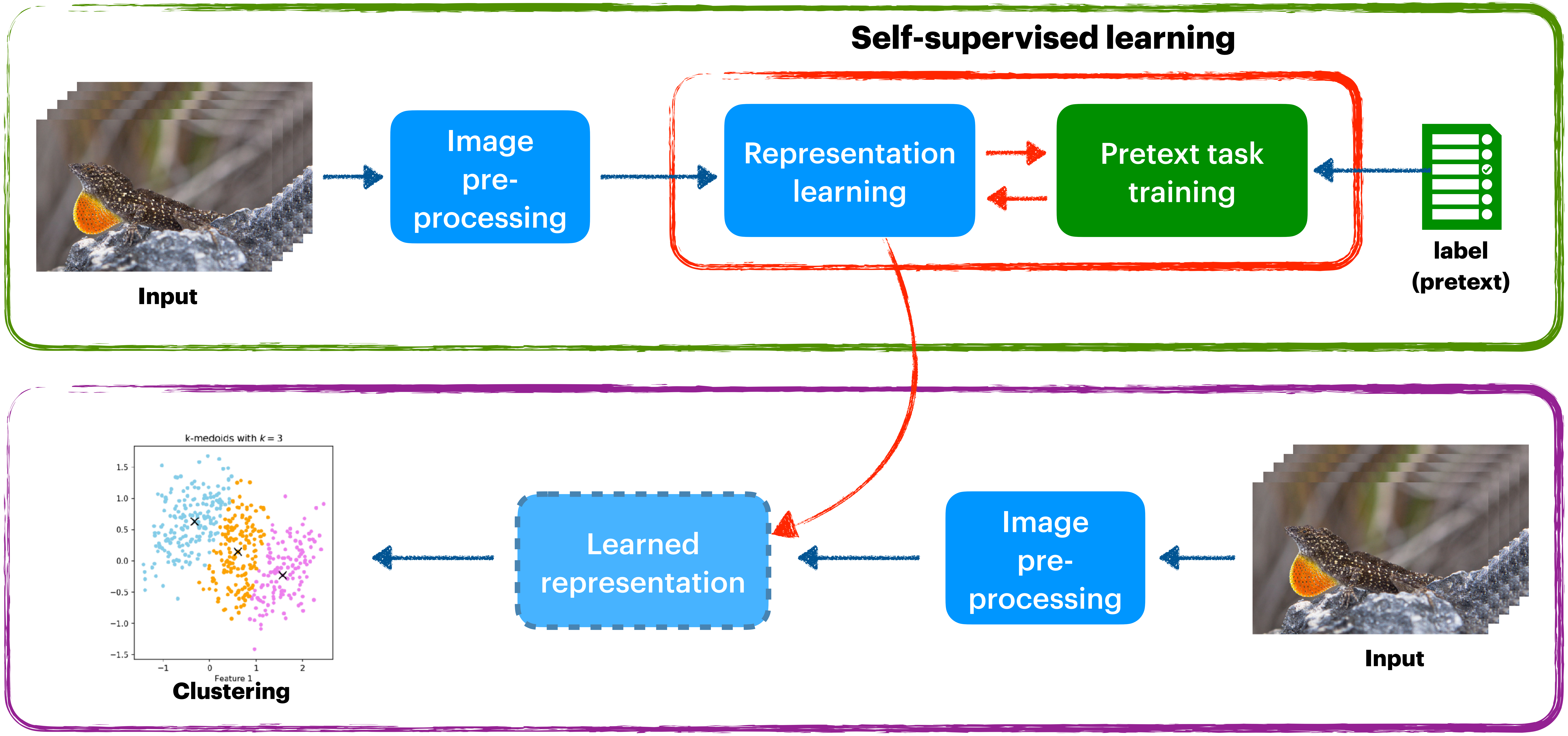
Choose the maximal activation (neuron)

Update parameters **only for maximal activation neuron**

Complete scheme



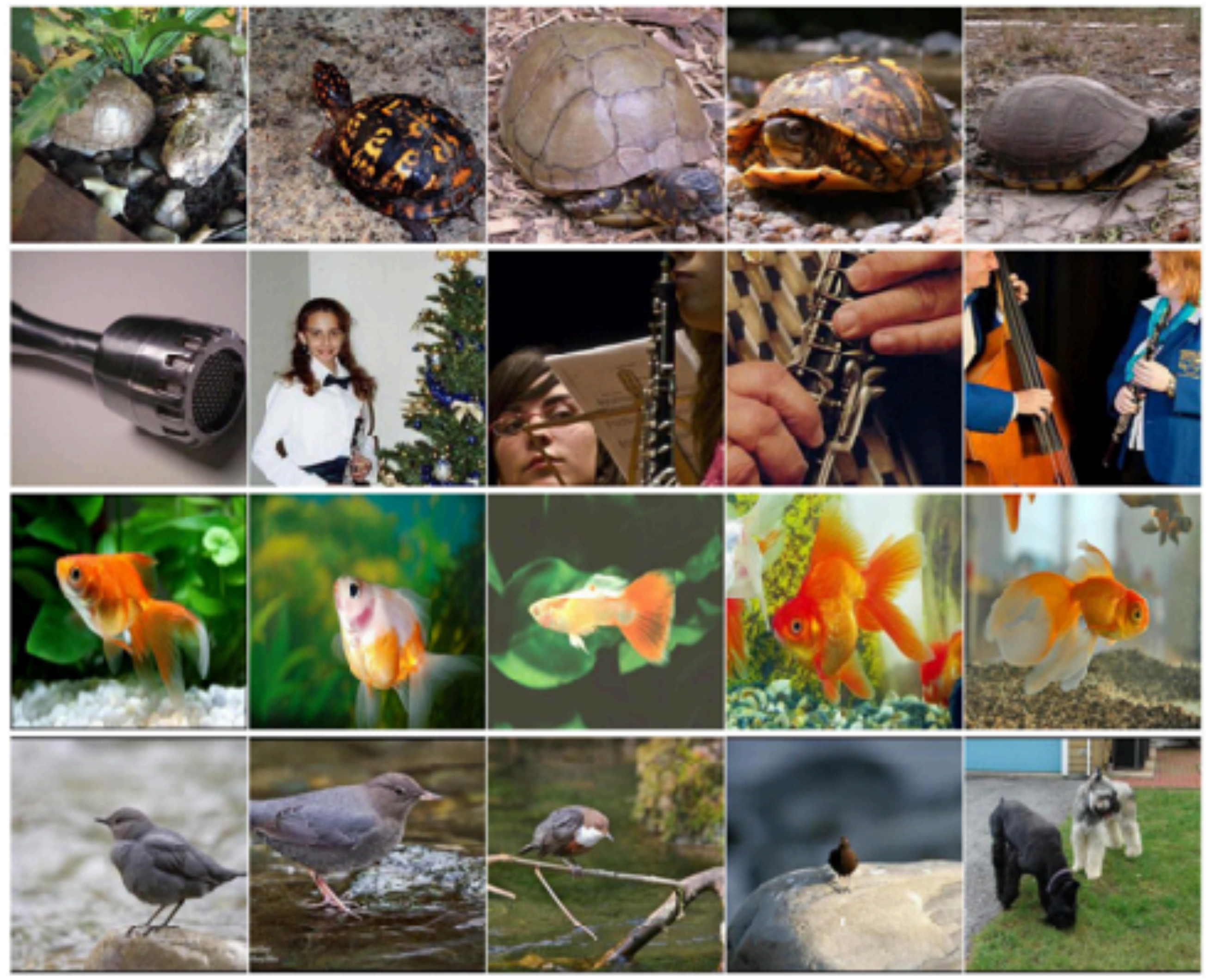
Complete scheme

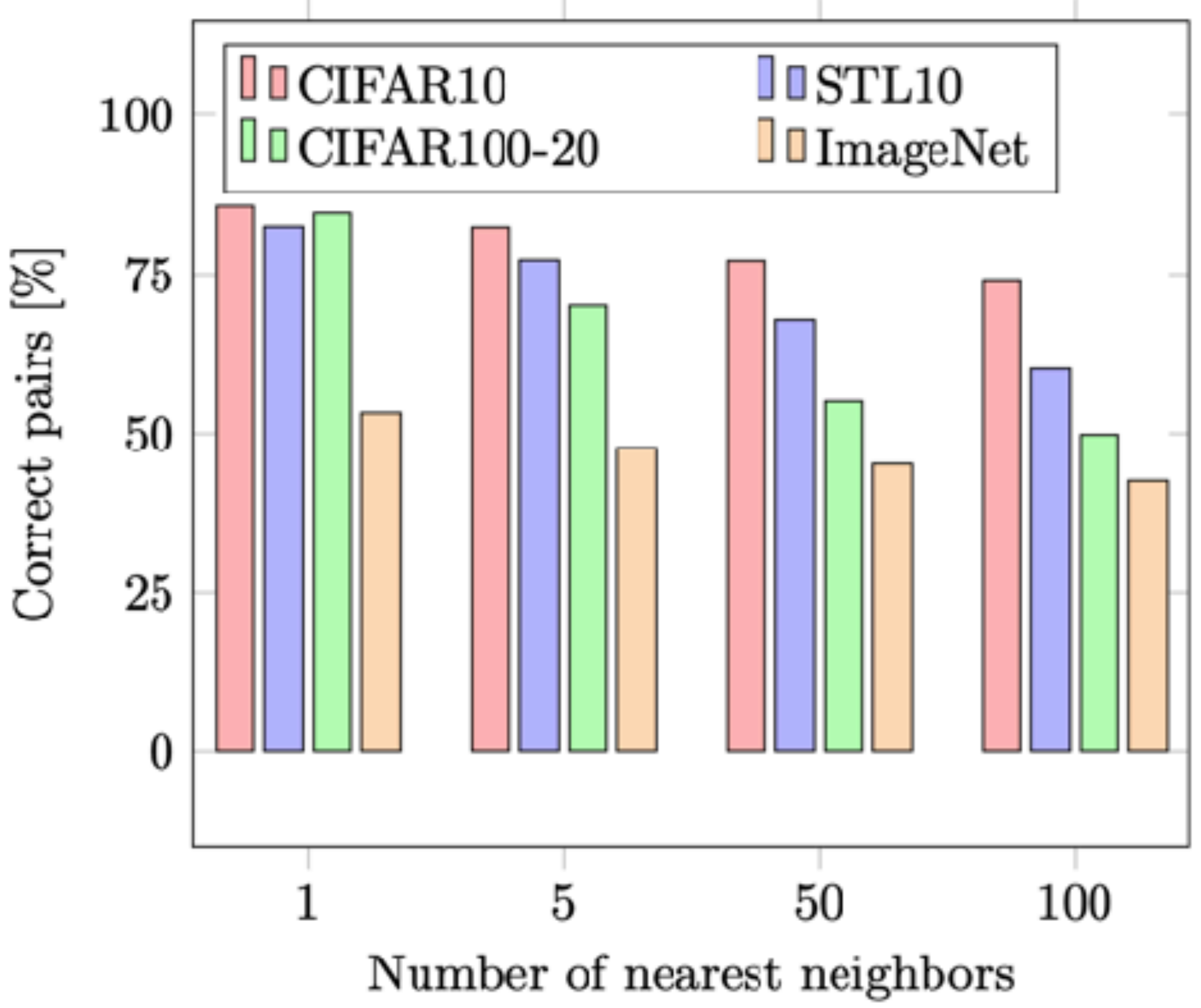


4. SCAN:Semantic Clustering by Adopting Nearest neighbors.

SCAN (2020)

- Combines self-supervised learning and clustering based approach in separated stages.
- Self-supervised for representation learning.
- k-Nearest Neighbors based loss for clusters detection.

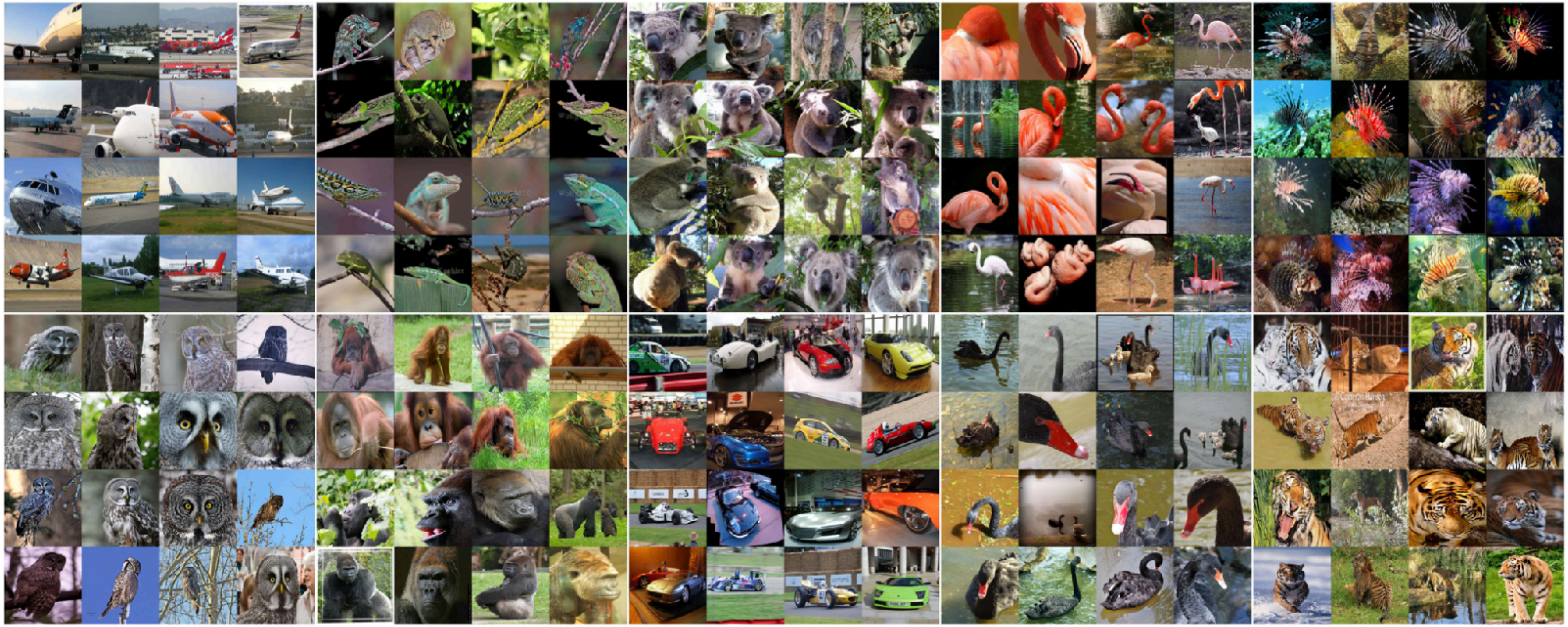




Method	Backbone	Labels	Top-1	Top-5
Supervised Baseline	ResNet-50	✓	25.4	48.4
Pseudo-Label	ResNet-50	✓	-	51.6
VAT + Entropy Min. [56]	ResNet-50	✓	-	47.0
InstDisc [51]	ResNet-50	✓	-	39.2
BigBiGAN [15]	ResNet-50(4x)	✓	-	55.2
PIRL [32]	ResNet-50	✓	-	57.2
CPC v2 [20]	ResNet-161	✓	52.7	77.9
SimCLR [7]	ResNet-50	✓	48.3	75.5
SCAN (Ours)	ResNet-50	✗	39.9	60.0

(Gansbeke et. al., 2020) SCAN: Learning to Classify Images without Labels.

SCAN (2020)



(Gansbeke et. al., 2020) SCAN: Learning to Classify Images without Labels.

5. Python Practice: Image Classification

Unsupervised image classification

[K-Means clustering for image classification]

6. Conclusions

Conclusions

- Self-supervised learning is like **a) supervised**: as it learns using a labelled data in a classification or regression task, and like **b) unsupervised**: as it does not requires the data to have real labels.
- Self-supervised learning is not unsupervised learning as it does not clusters or groups data.
- Clustering techniques can be used in order to obtain pseudo-labels from unlabelled data.
- Unsupervised image classification is a relative new field which combines the self-supervised learning for feature/representation learning, and unsupervised learning in order to label data.