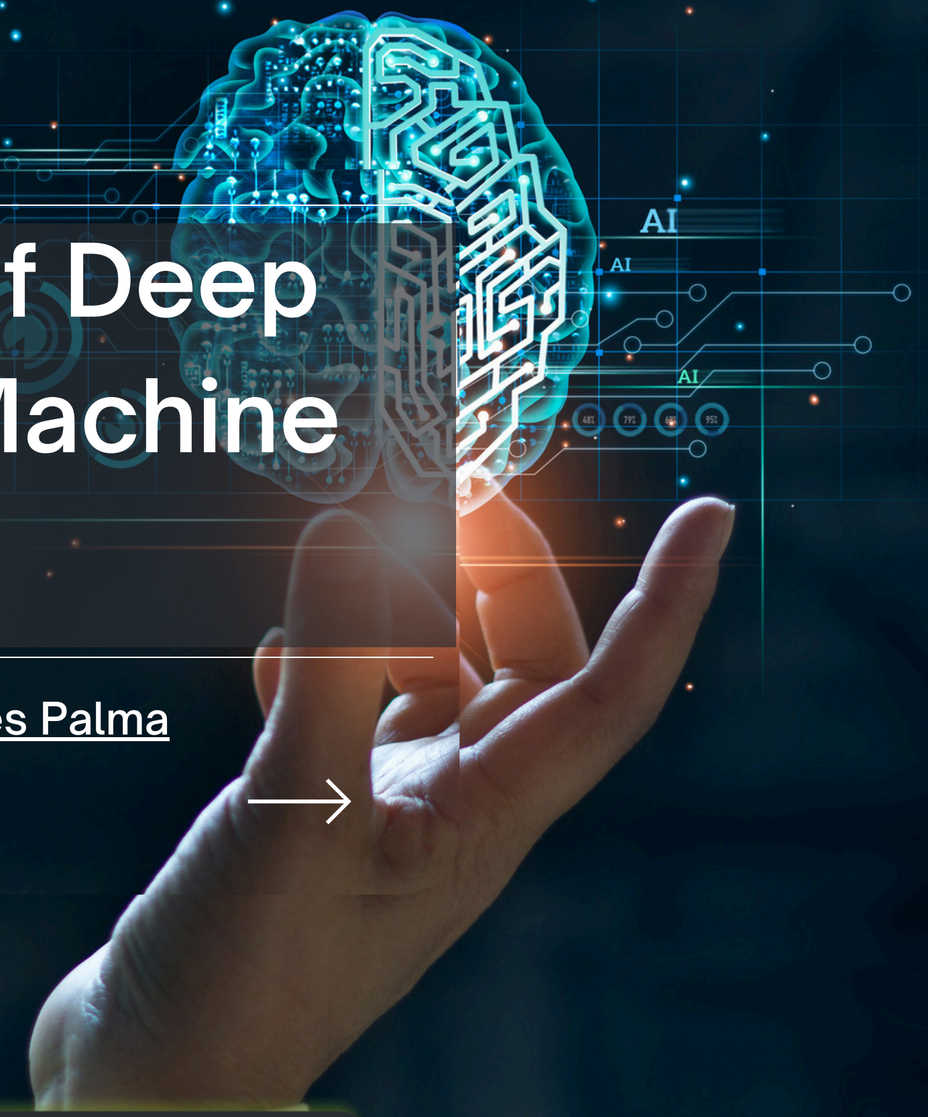
Applications of Deep Learning and Machine Vision

Presented by Gabriel Rodrigues Palma



ESALQ

Course introduction

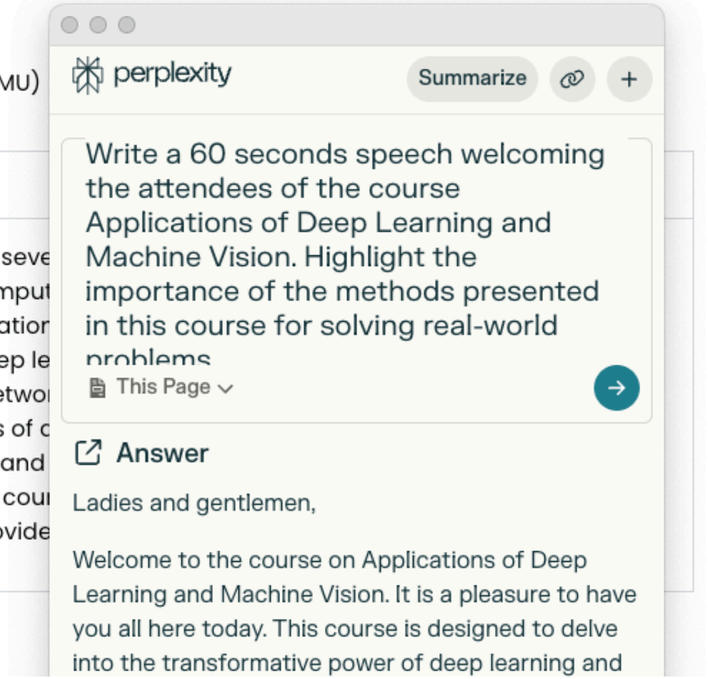
Minicurso 1 - MC1 – 29/05 e 30/05 (08:00 – 10:00) – Sala 314

Título: Applications of Deep Learning and Machine Vision

Ministrantes: Gabriel R. Palma e Darshana Jayakumari – Maynooth University (MU)

— Resumo:

Machine vision has produced many helpful image-processing techniques in seve
segmentation. Machine vision is an interdisciplinary discipline combining comput
deep learning to solve vision problems. Common problems, such as classificatior
these research fields. These techniques have applications in many areas. Deep le
classification, focusing on deep neural networks and Convolutional Neural Netwo
applied to image classification. This course aims to introduce basic concepts of c
classification using CNNs. To illustrate these methods a dataset of medically and
programming language and the package Tensorflow will be used during the cou
language. All the course materials and additional online resources will be provide



Course introduction

Applications of Deep Learning and Machine Vision

ESALQ

Overview

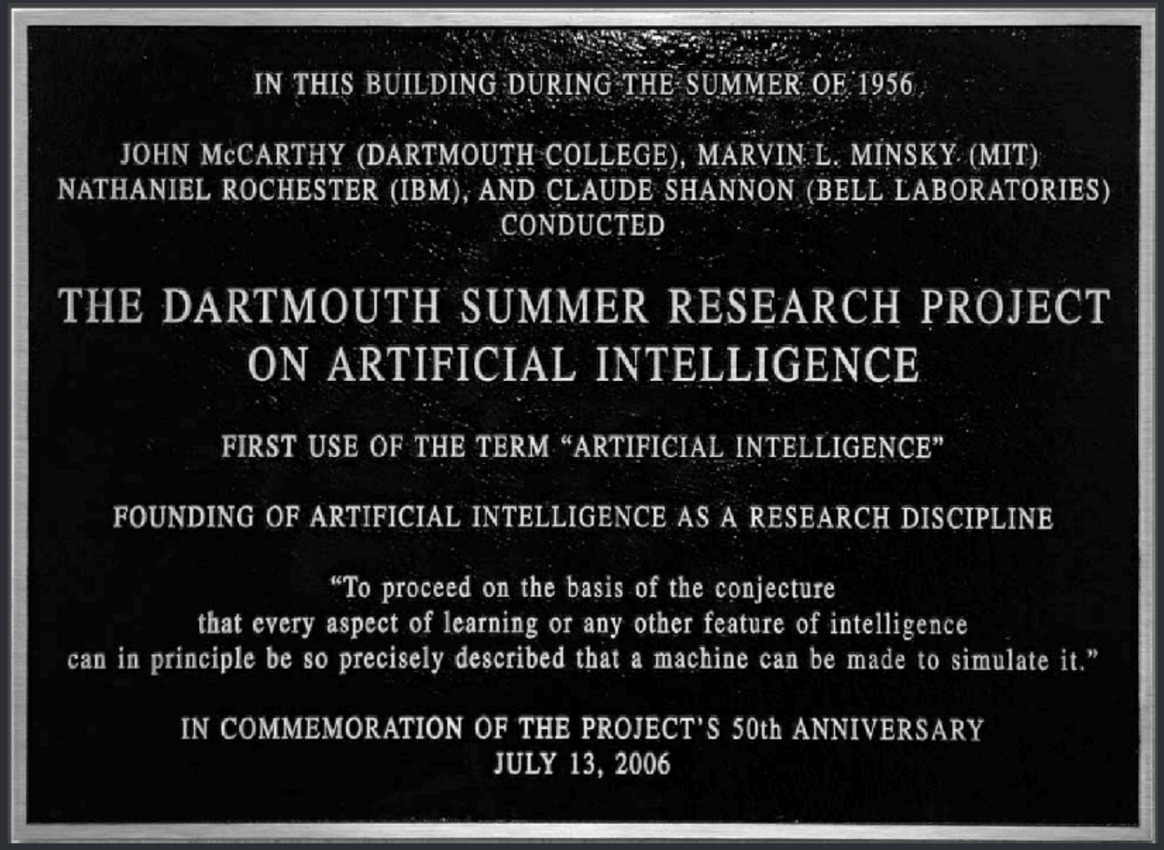
Day 1

- An introduction to Machine Vision:
  - Overview of a machine vision pipeline workflow;
  - An introduction to OpenCV in Python;
- An introduction to Artificial Neural Networks:
  - What is a Deep Neural Network?
  - DNN using TensorFlow

Day 2

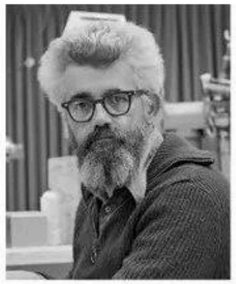
- An introduction to Convolutional Neural Networks;
  - What are the basic operations of a CNN?
  - How do we train a CNN?
- An introduction to Transfer Learning;
  - How can we take advantage of transfer learning?

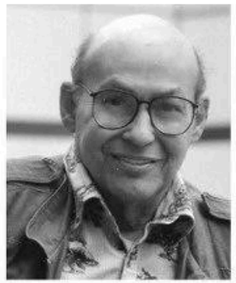
An introduction to Machine Vision

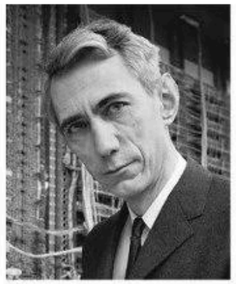


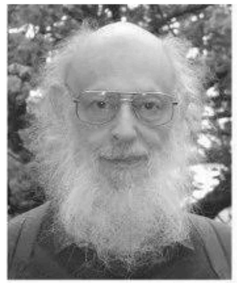
An introduction to Machine Vision

1956 Dartmouth Conference: The Founding Fathers of AI


John MacCarthy


Marvin Minsky

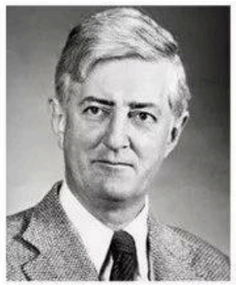

Claude Shannon


Ray Solomonoff


Alan Newell


Herbert Simon

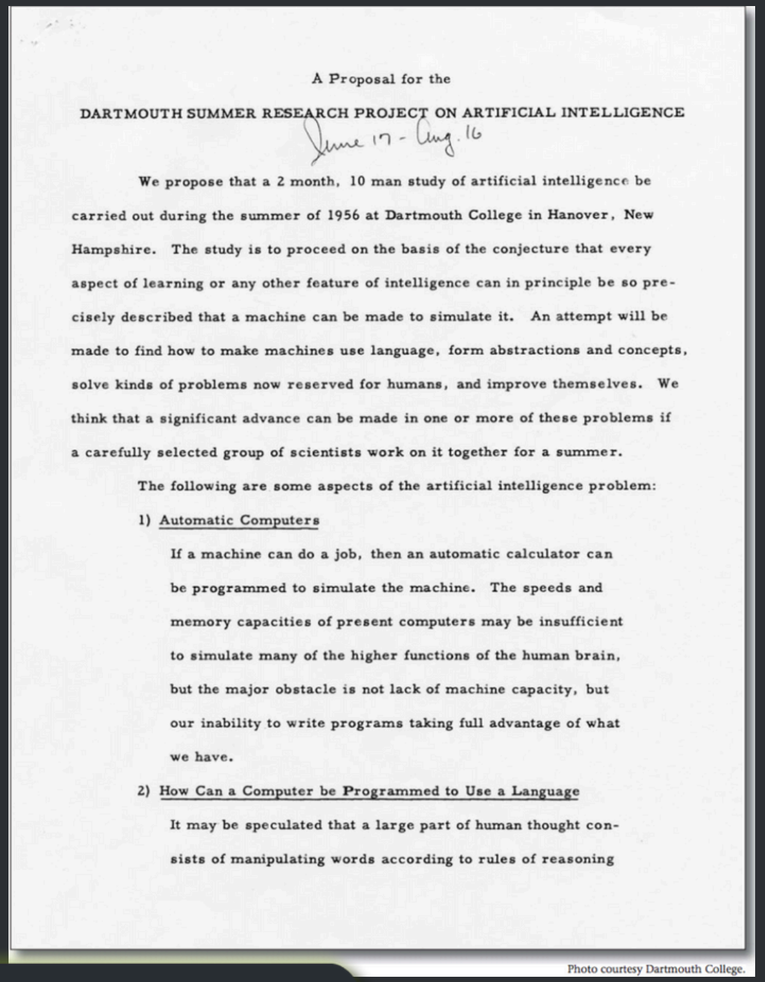

Arthur Samuel


Oliver Selfridge


Nathaniel Rochester


Trenchard More

An introduction to Machine Vision



A Proposal for the

DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE

June 17 - Aug. 16

We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer.

The following are some aspects of the artificial intelligence problem:

1) Automatic Computers

If a machine can do a job, then an automatic calculator can be programmed to simulate the machine. The speeds and memory capacities of present computers may be insufficient to simulate many of the higher functions of the human brain, but the major obstacle is not lack of machine capacity, but our inability to write programs taking full advantage of what we have.

2) How Can a Computer be Programmed to Use a Language

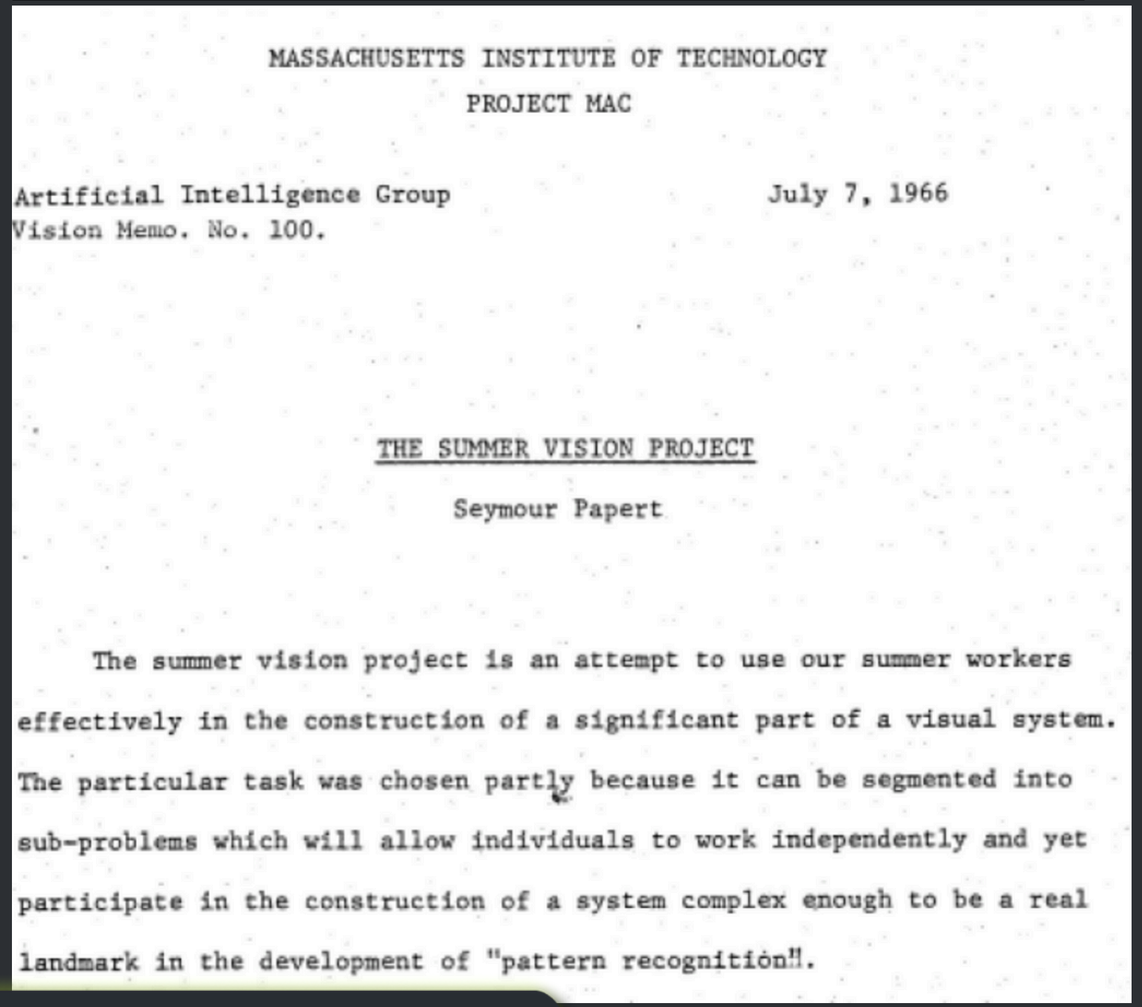
It may be speculated that a large part of human thought consists of manipulating words according to rules of reasoning

Photo courtesy Dartmouth College.

Applications of Deep Learning and Machine Vision

ESALQ

An introduction to Machine Vision



MASSACHUSETTS INSTITUTE OF TECHNOLOGY

PROJECT MAC

Artificial Intelligence Group — July 7, 1966
Vision Memo. No. 100.

THE SUMMER VISION PROJECT

Seymour Papert

The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".

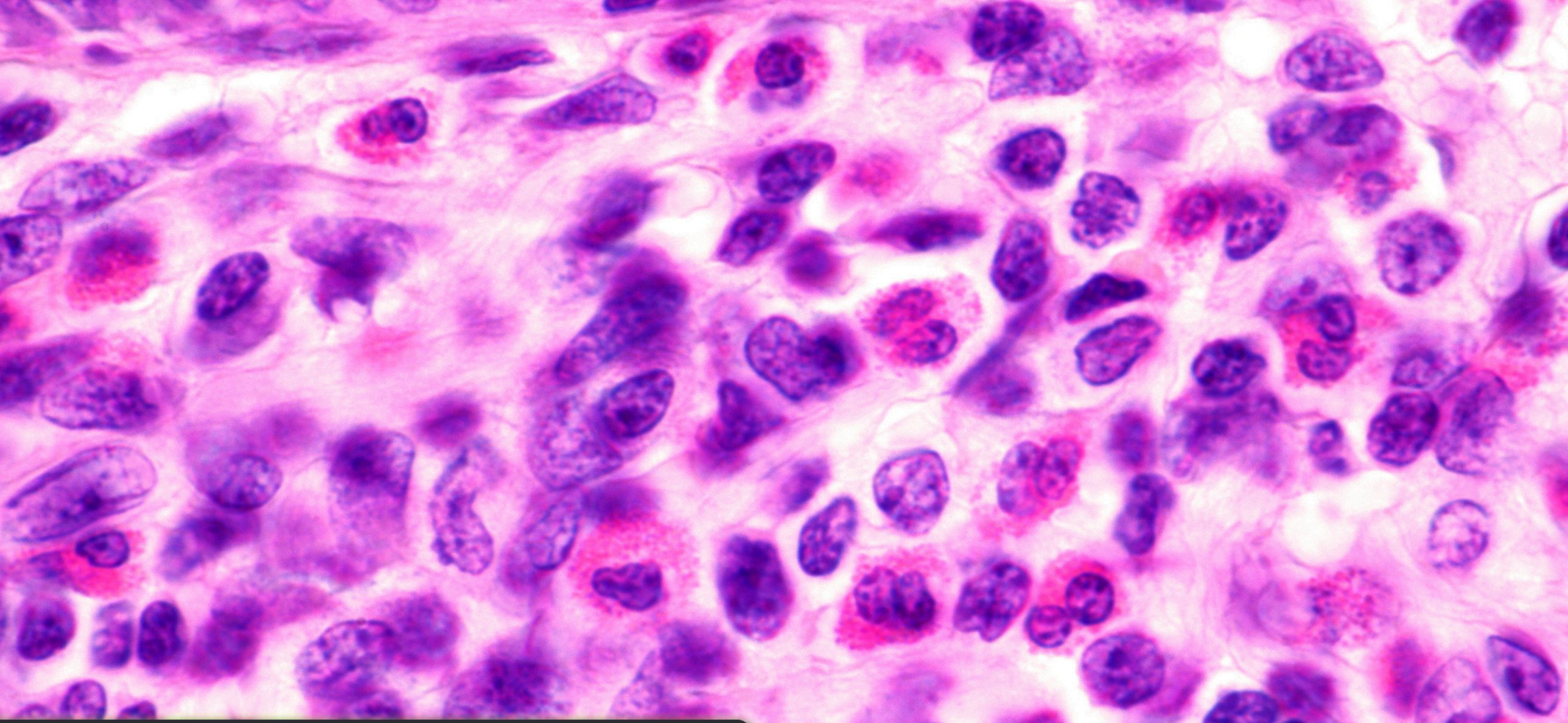
ESALQ

Applications of Deep Learning and Machine Vision

ESALQ

Applications of Deep Learning and Machine Vision

ESALQ

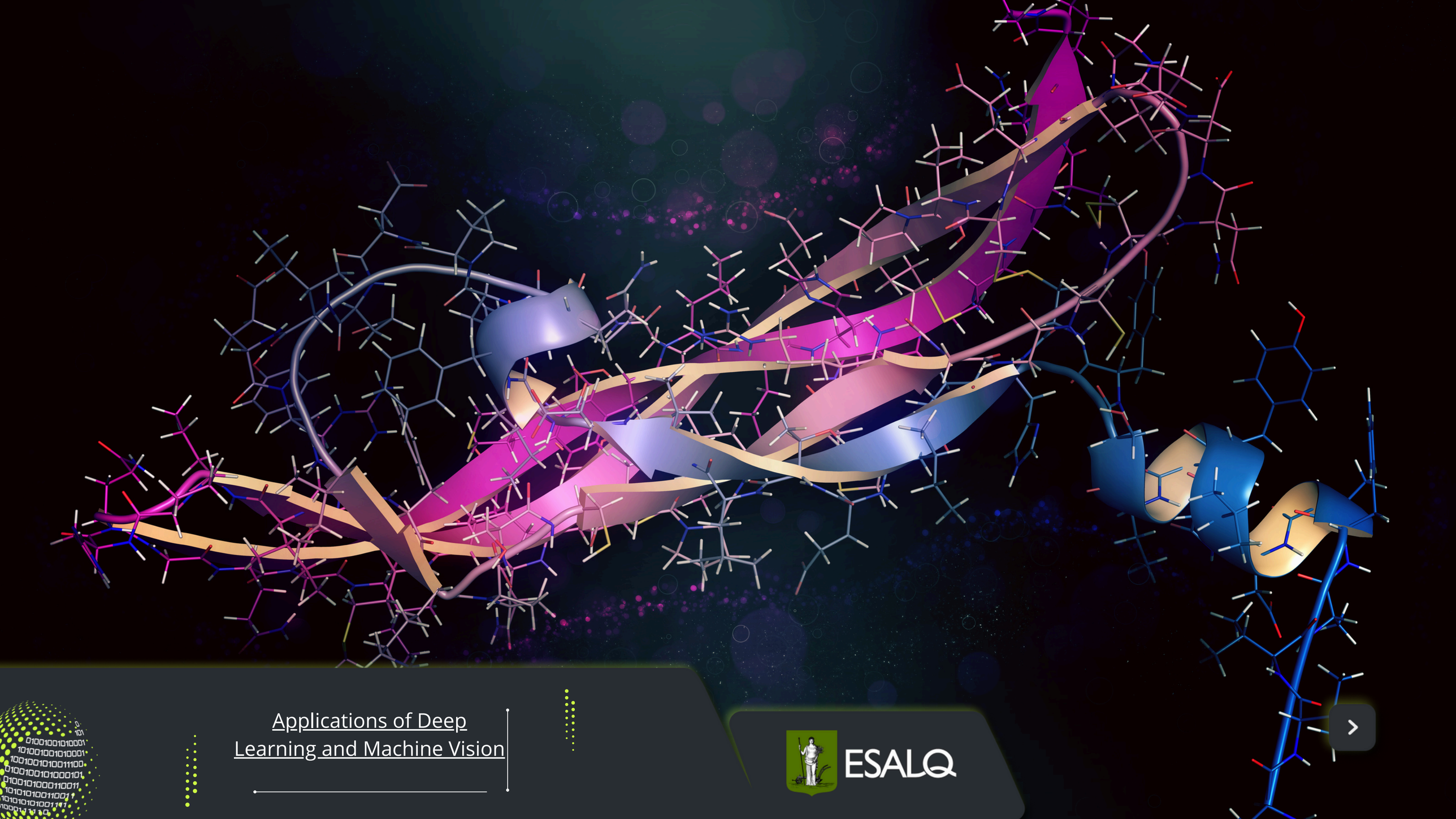


Applications of Deep Learning and Machine Vision

ESALQ

Applications of Deep Learning and Machine Vision

ESALQ



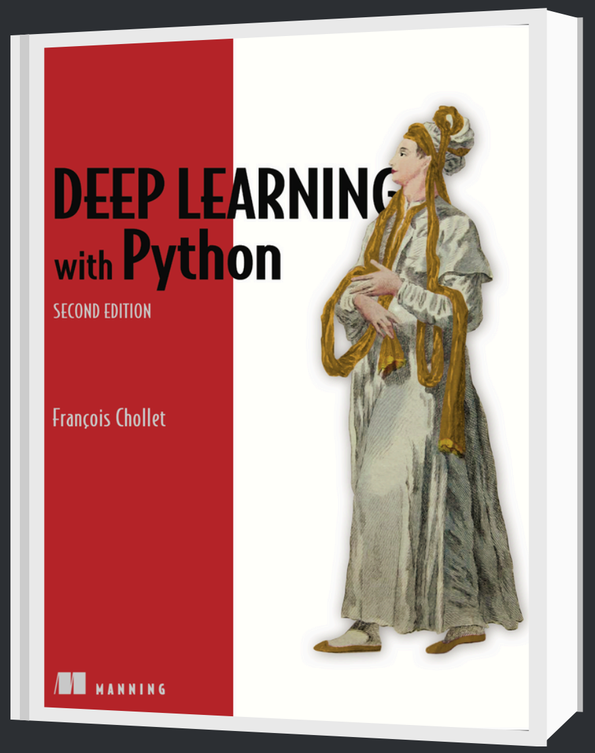
Applications of Deep Learning and Machine Vision

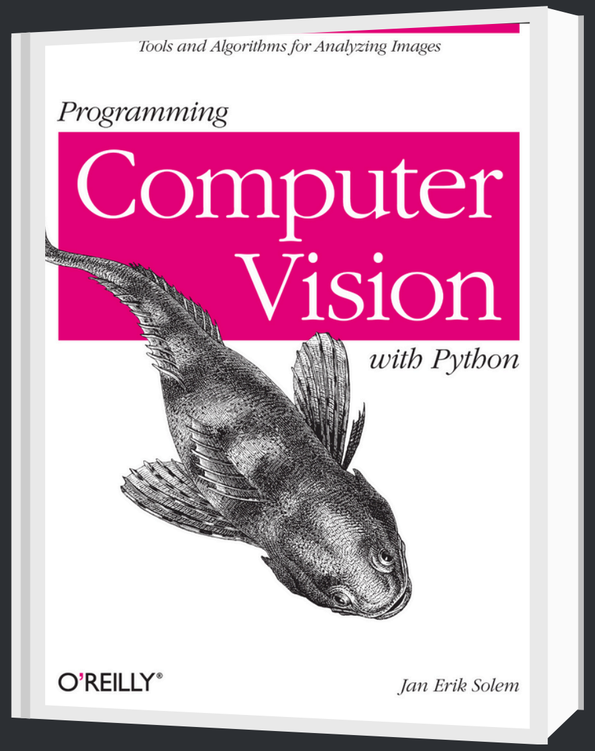
ESALQ

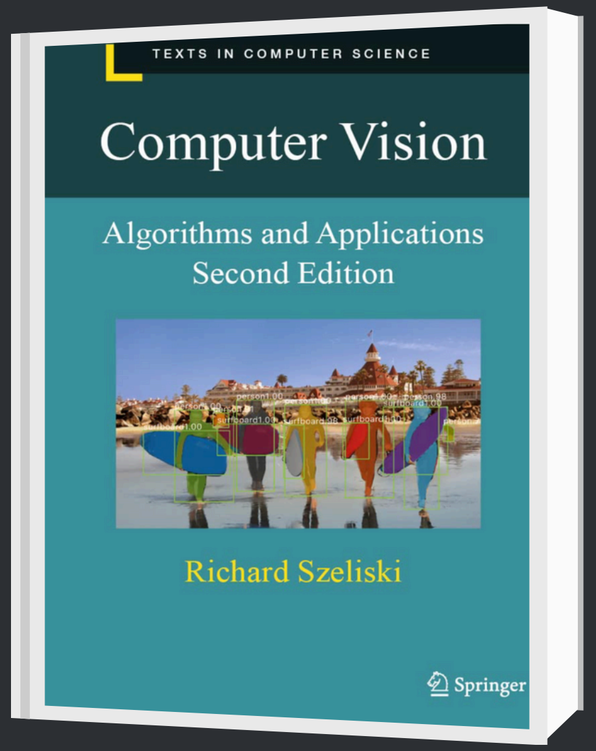
Applications of Deep Learning and Machine Vision

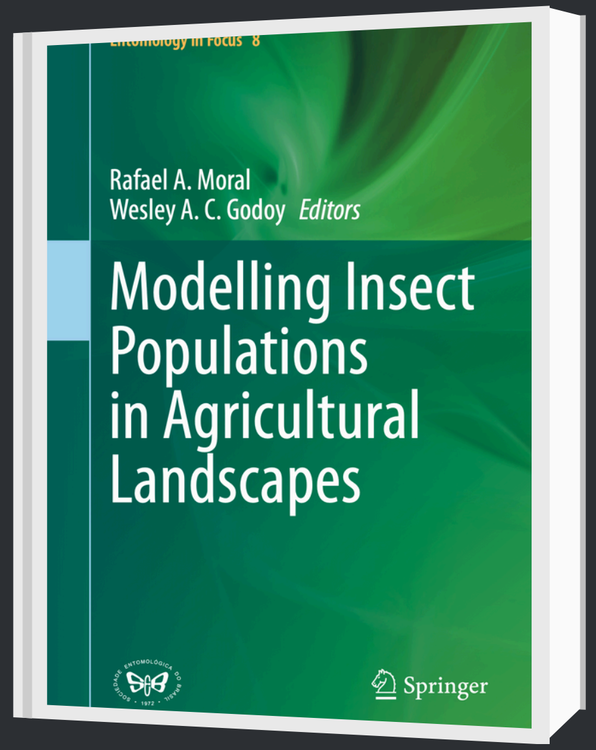
ESALQ

An introduction to Machine Vision









Applications of Deep Learning and Machine Vision

ESALQ

An introduction to Machine Vision

Chapter 9
Machine Vision Applied to Entomology

Gabriel R. Palma, Conor P. Hackett, and Charles Markham

Abstract Machine vision has produced many helpful image processing techniques in several fields, such as object detection, classification, and segmentation. Machine vision is an interdisciplinary field that combines methods from computer vision and machine learning, mainly deep learning to solve issues. Common issues, such as classification and localisation, are typical examples that combine these research fields. These techniques have many applications in entomology. In the entomological context, classification methods identify insect species or features of an insect, and localisation methods return their position within a scene. The available techniques for localisation and classification benefit the entomological community by assisting specialists. This chapter introduces the application of machine vision methods to identify and localise insects and the features of their anatomy. The machine vision pipeline, image descriptors, and fundamental methods, such as thresholding, blob, and contour detection, are presented. Deep Learning (DL) methods are applied for insect classification, focusing on Deep Neural Networks and Convolutional Neural Networks (CNNs), including concepts of transfer learning applied to insect classification. A dataset of medically and forensically important flies, Diptera, is used to illustrate these methods. CNN-based methods are described for insect localisation, and semantic segmentation using the U-Net architecture trained on the dataset of flies is presented. The dimensional reduction method PaCMAP is applied to visualise the feature extracted from the fly datasets. The Grad-CAM method is used on the flies datasets to assist insect object localisation. Finally, different platforms of computing are presented and compared under three different metrics of performance, power consumption, and operating temperature.

Keywords Artificial intelligence · Computer vision · Deep learning · Image classification · Machine learning

Taxonomic classification is one of the pillars required for implementing Integrated Pest Management (IPM) methods (Jaleel et al., 2018; Deguine et al., 2021; Shimbori1 et al., 2023). It allows the identification of important pests with a detailed description of their biology. Researchers use the biological information gathered by entomologists to create management theories, yielding essential technologies to control the insect pest population (Stenberg et al., 2021). These new findings positively impact our society, for example, the economic gains that biological control has offered to agriculture. These gains can reduce environmental impact and cost and enhance food security (Bale et al., 2008; van Wilgen et al., 2020; Stenberg et al., 2021).

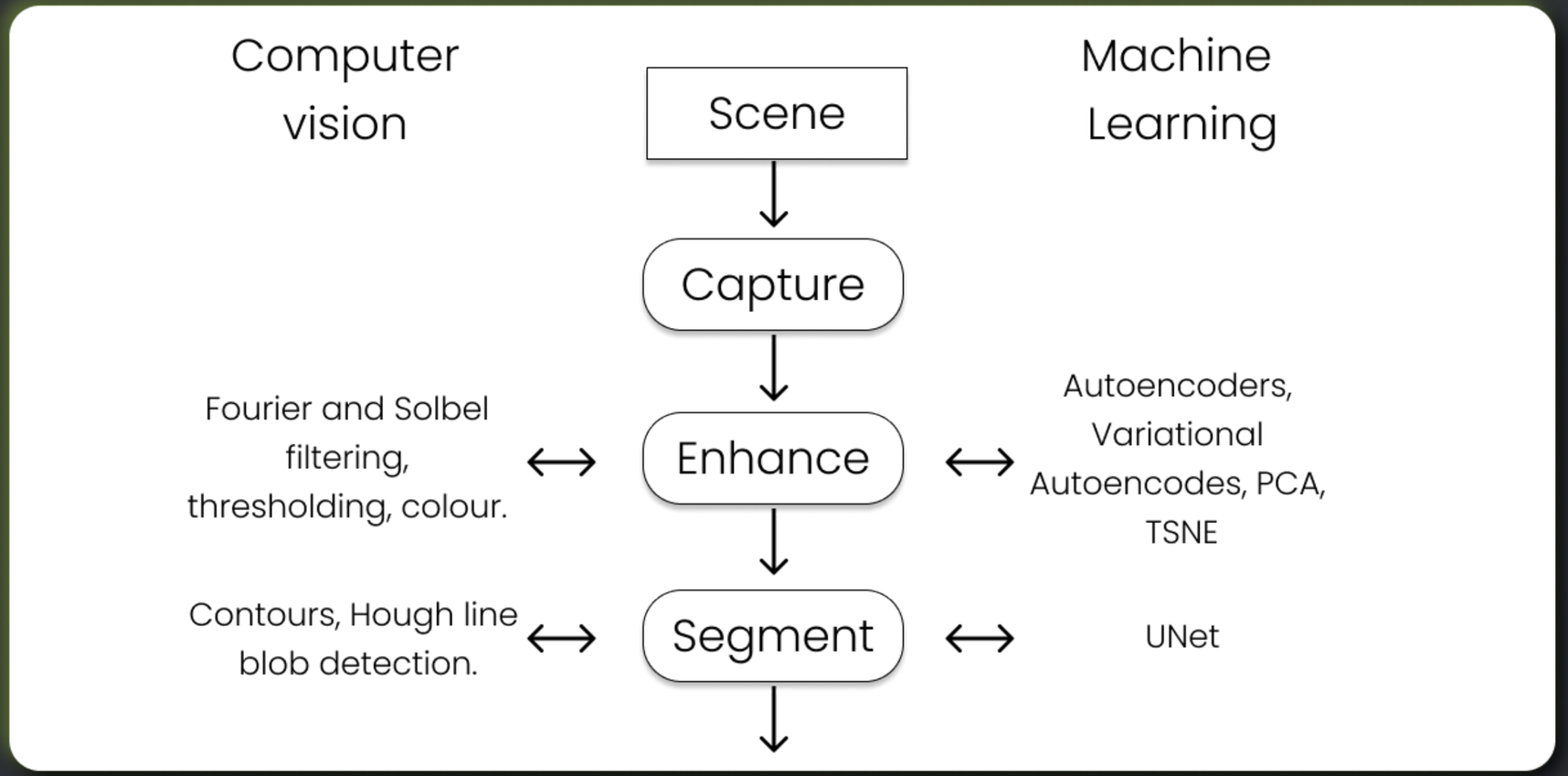
Entomological research has provided direct contributions to the control of pest management in agriculture, such as the introduction of the species *Tamarixia radiata* that allowed the biological control of the pest *Diaphorina citri* (Étienne et al., 2001; Qureshi et al., 2009; Chow & Sétamou, 2022). The description and identification of parasitoid species *Trichogramma galloi* (Zucchi, 1988) acted as an important insect for the management of *Diatraea saccharalis* (Cônsoli & Parra, 1996; Camarozano et al., 2021). These studies also present the spatial distribution of the targeted species and recommendations about the parasitism rate. The biological descriptions allow further research and quantitative approaches, such as mathematical, statistical, and computational modelling (Camarozano et al., 2021).

In addition, entomological research offers vital information to wildlife conservation by gathering awareness of global biodiversity and providing evidence of collected species with clear documentation in national museums as evidence. As this is such an important field, many researchers have tried to develop techniques to automate insect identification. By combining the collected knowledge from computer vision, deep learning, and labelled datasets created by entomologists, it is now possible to create algorithms that can automate or assist the classification of insects in different taxonomical levels, such as order (Ozdemir & Kunduraci, 2022), genus (Ong & Ahmad, 2022), and species (Thenmozhi & Reddy, 2019).

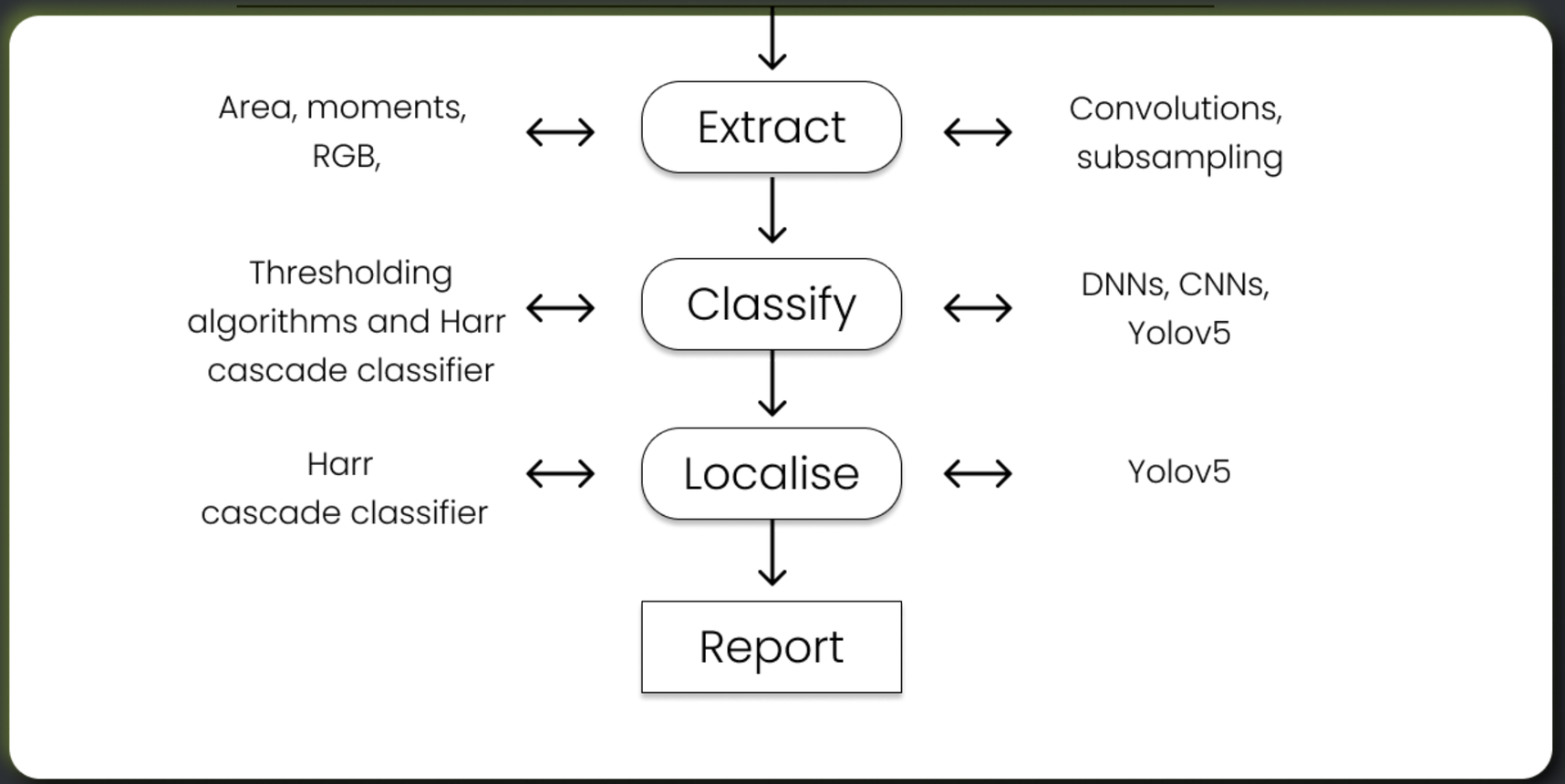
Several researchers have implemented computer vision techniques to automate insect identification of species such as *Bemisia tabaci*, *Sesamia inferens*, and *Chilo suppressalis* (Zayas & Flinn, 1998; Zhigang et al., 2003; Larios et al., 2008; Solis-Sánchez et al., 2009; Yang et al., 2010; Asefpour Vakilian & Massah, 2013; Favret & Sieracki, 2016; Qing et al., 2020; Kasinathan & Uyyala, 2021; Kasinathan et al., 2021). These results have shown promising results by obtaining features and applying machine learning algorithms, such as Support Vector Machines (SVMs), Deep Neural Networks (DNNs), Linear Discriminant Analysis (LDA), and Random Forest (RF) (Hastie et al., 2004; Goodfellow et al., 2016; Mello & Ponti, 2018; Wani et al., 2020) to classify the species. These results illustrate the feasibility of combining machine learning and vision methods.

The introduction of deep Learning methods based on Deep Neural Networks has expanded machine vision applications

An introduction to Machine Vision



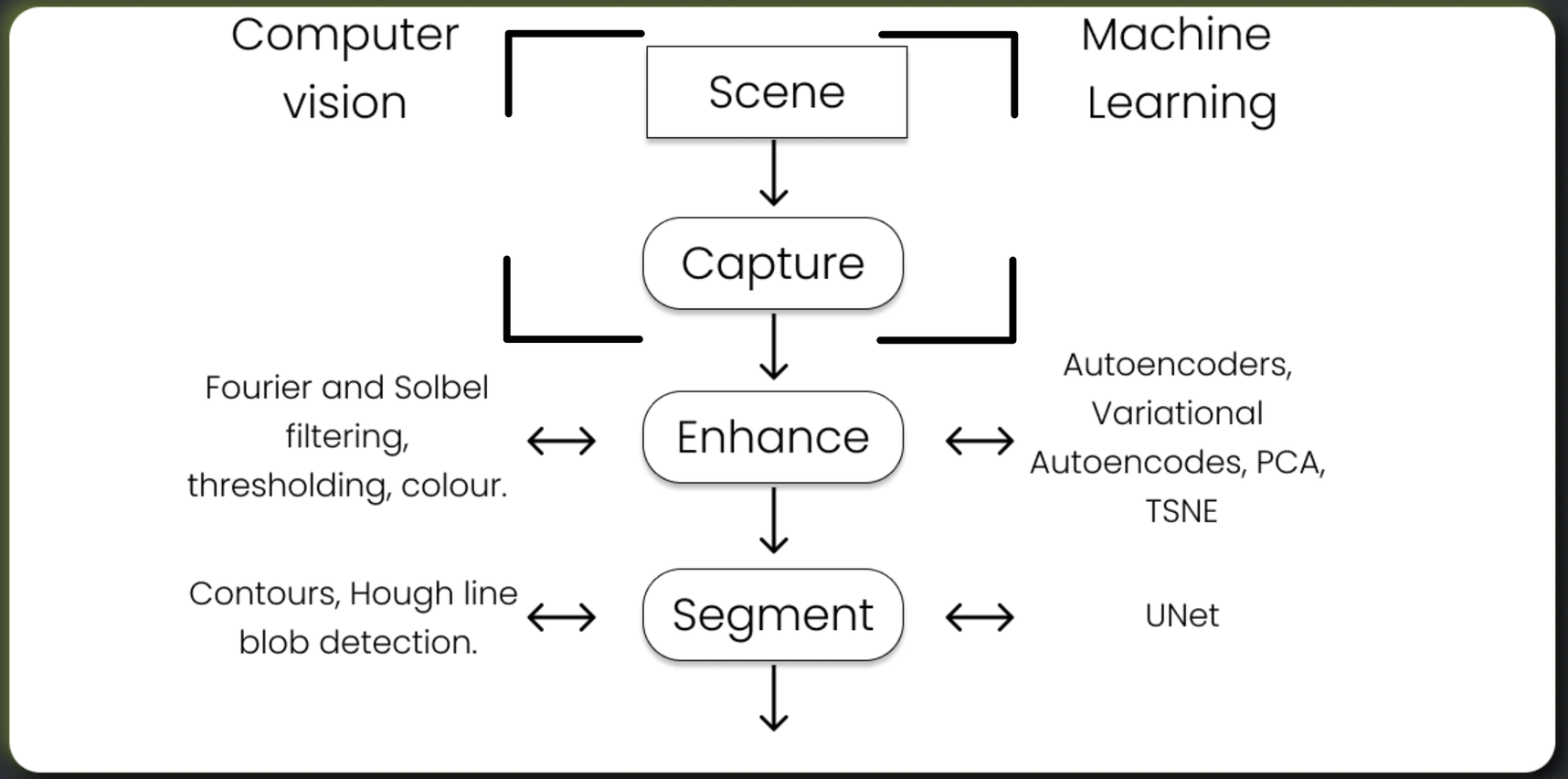
An introduction to Machine Vision



Applications of Deep Learning and Machine Vision

ESALQ

Overview of a machine vision pipeline workflow



Overview of a machine vision pipeline workflow

Stomorhina

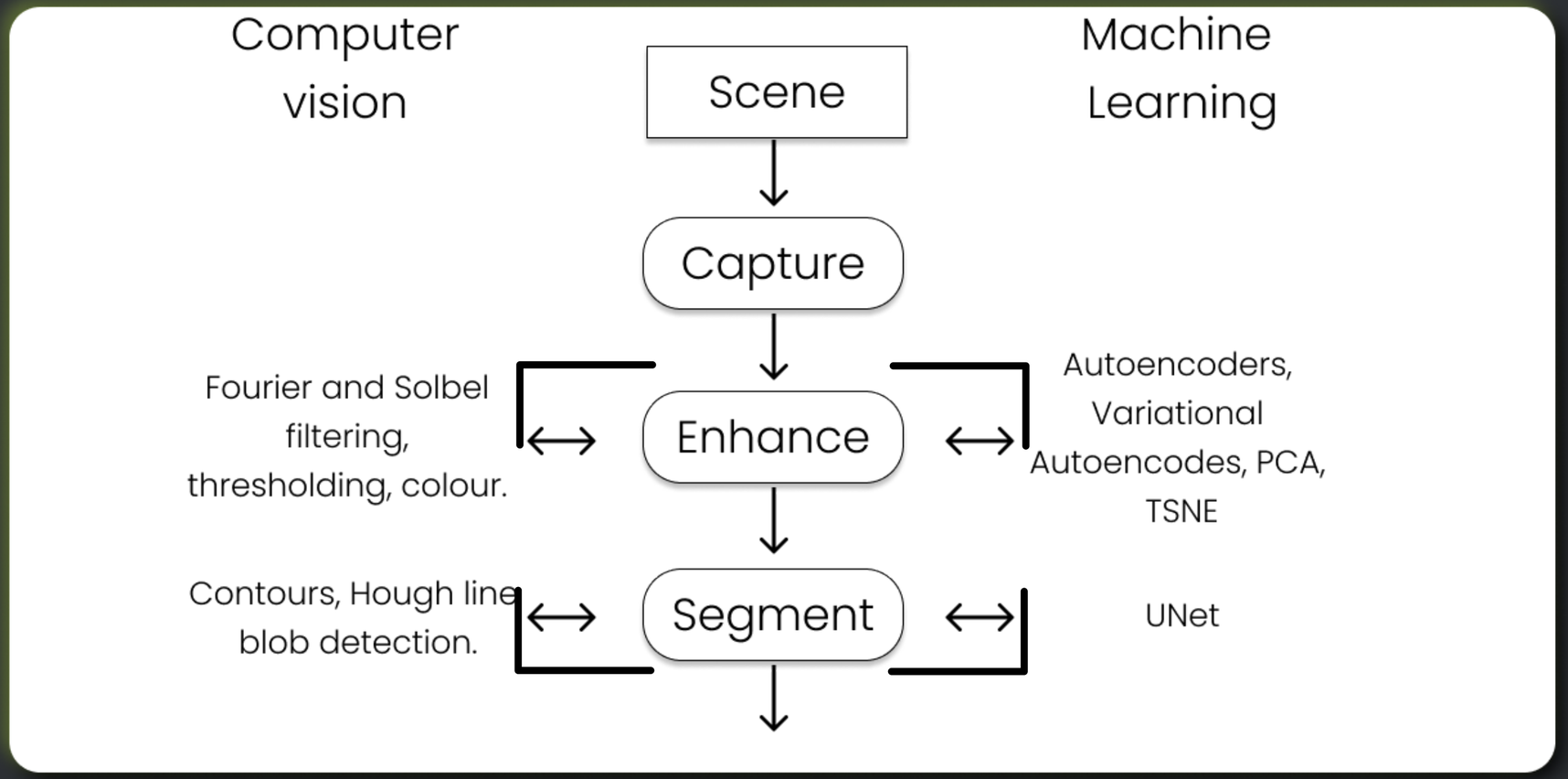
Sarcophaga

Rhiniinae

Lucilia

Chrysomya

Overview of a machine vision pipeline workflow

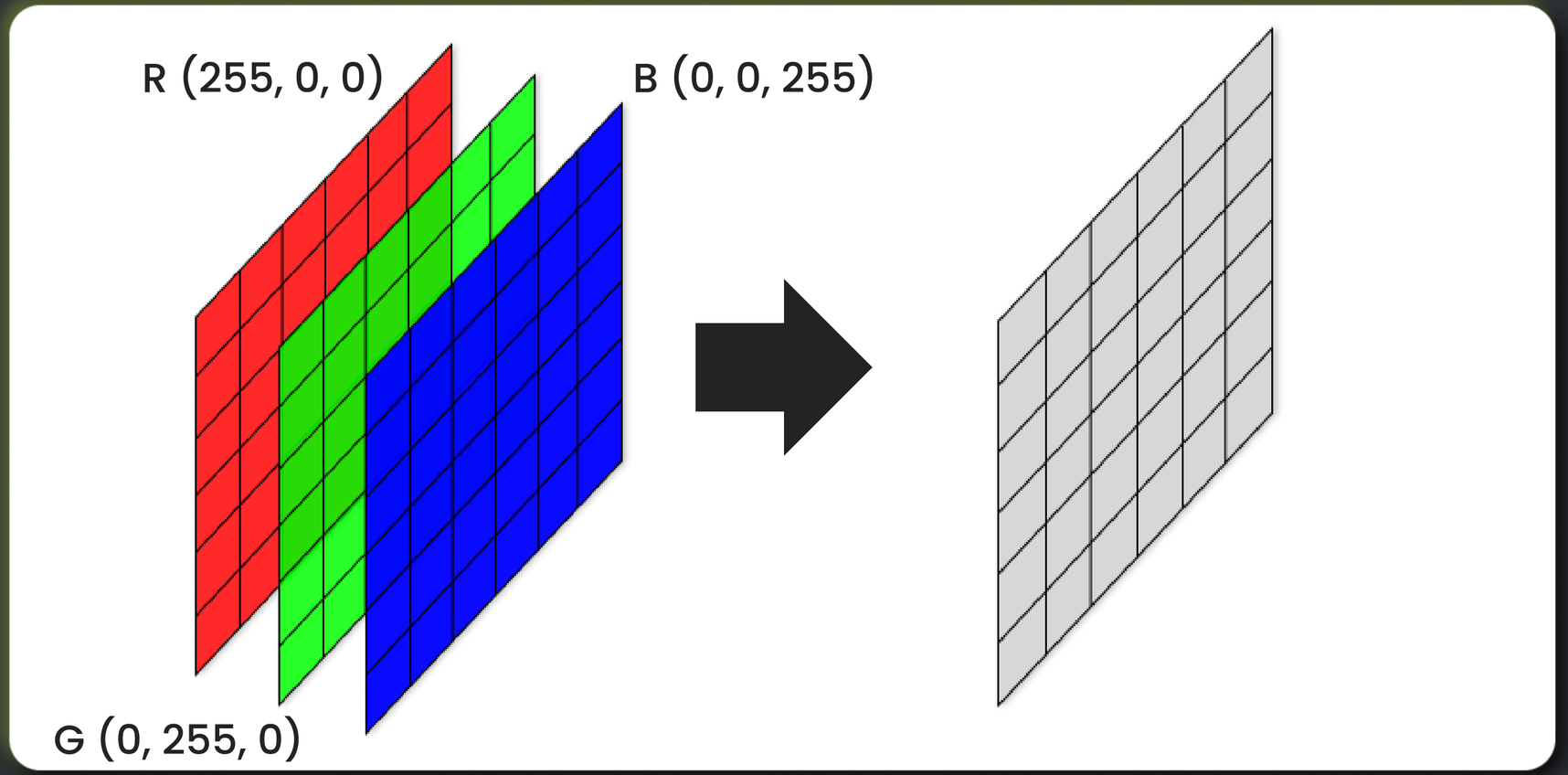


Overview of a machine vision pipeline workflow

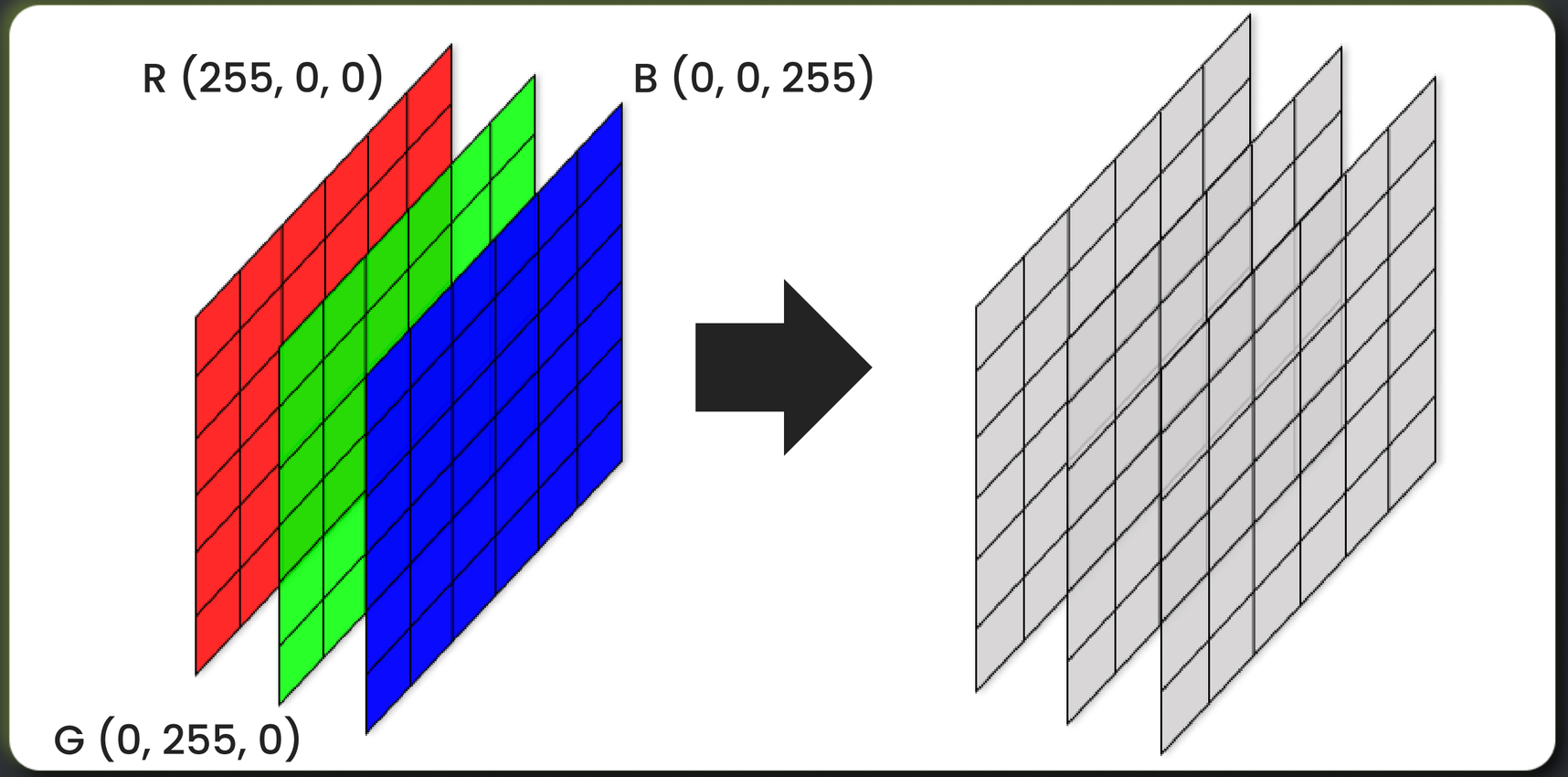
Enhance: Enhancement is a pre-processing step that modifies an image to assist with the segmentation. Modifications to an image made by enhancement include:

- Conversion of colour to greyscale (a black and white image)
- Edge detection;
- Filtering;
- Thresholding (identifying features of an image using their pixel value).

Overview of a machine vision pipeline workflow

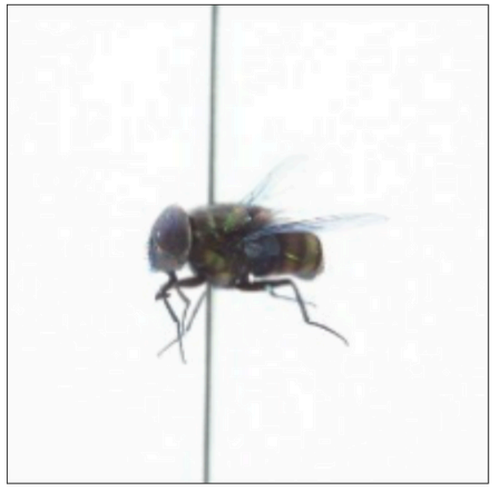


Overview of a machine vision pipeline workflow

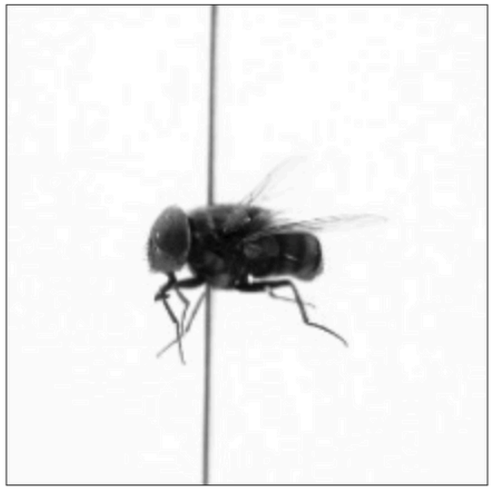


Overview of a machine vision pipeline workflow

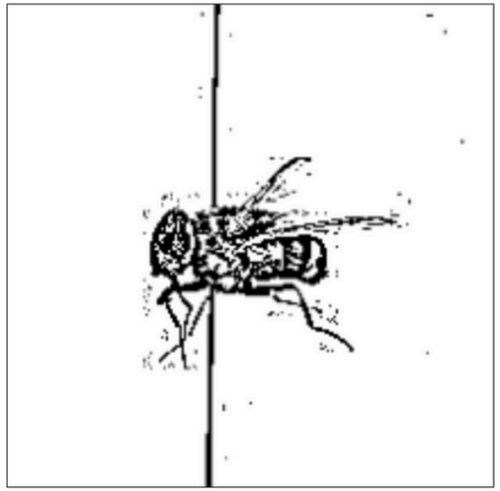
Colored image



Grayscale image



Thresholded image

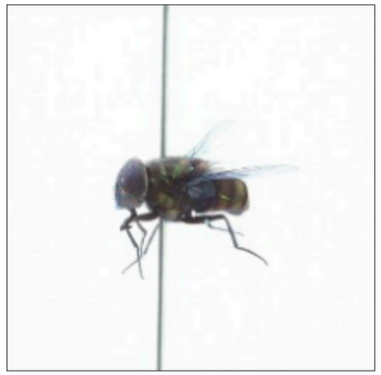
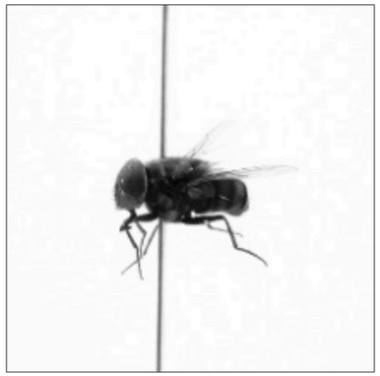
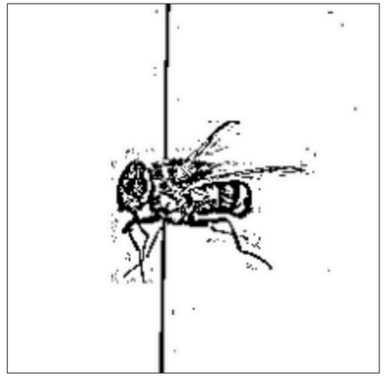
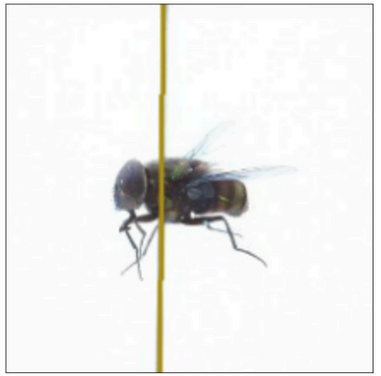


Overview of a machine vision pipeline workflow

Segment: Segmentation takes an image as input and returns a list of image features. The image features include blobs, lines, and contours.

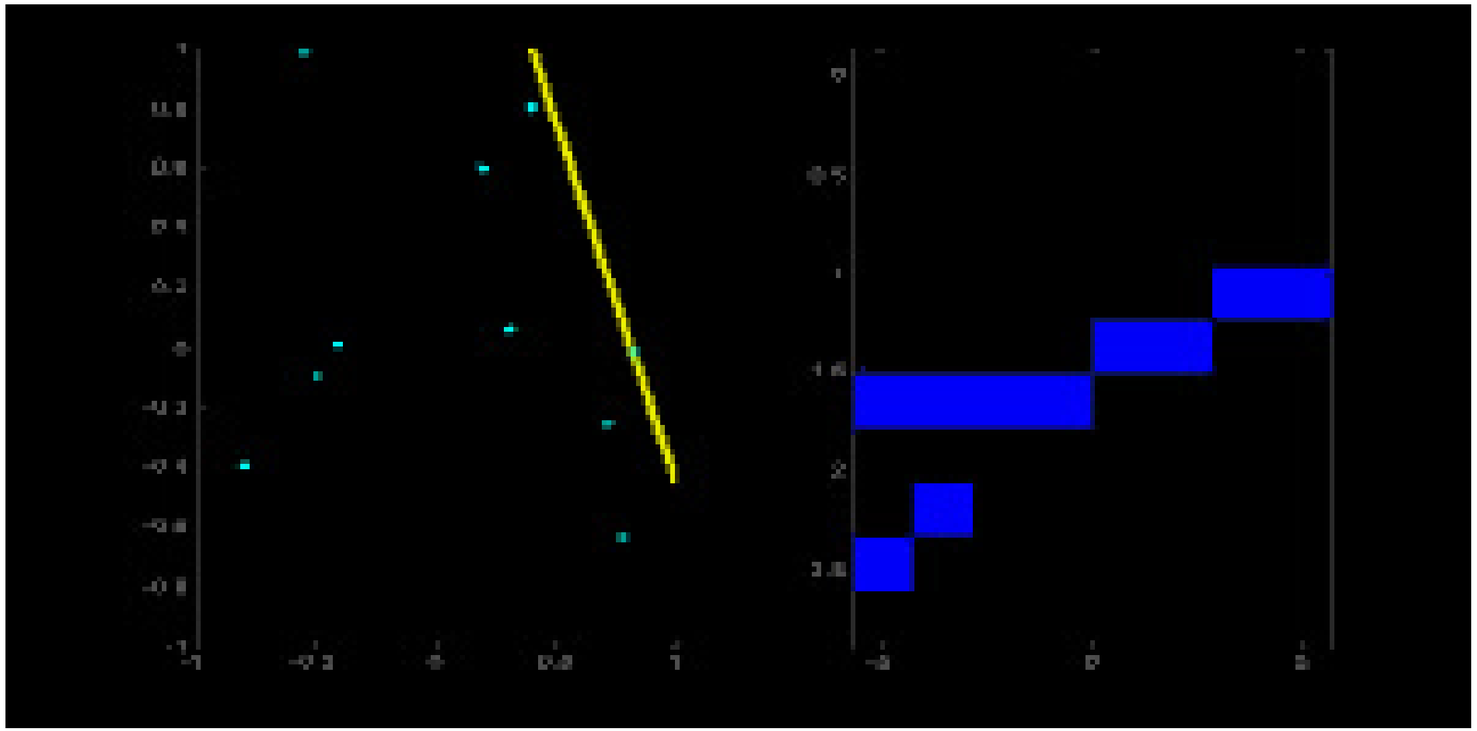
- A Hough transform identifies pixels that belong to lines in the image;
- A blob is a connected set of pixels in a thresholded image;
- The contour describes the bounding pixels of a segmented region.

Overview of a machine vision pipeline workflow

Colored image	Grayscale image	Thresholded image	Hough Line
			

Overview of a machine vision pipeline workflow

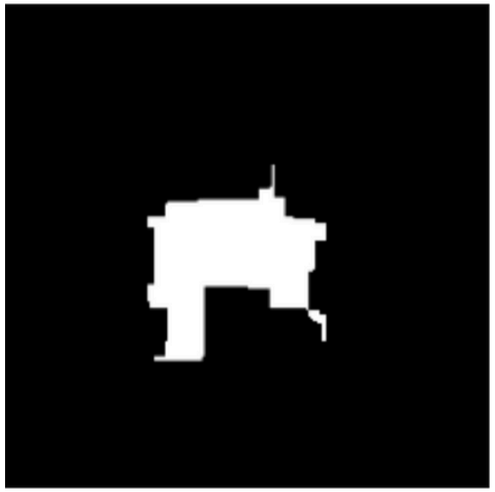
A Hough transform



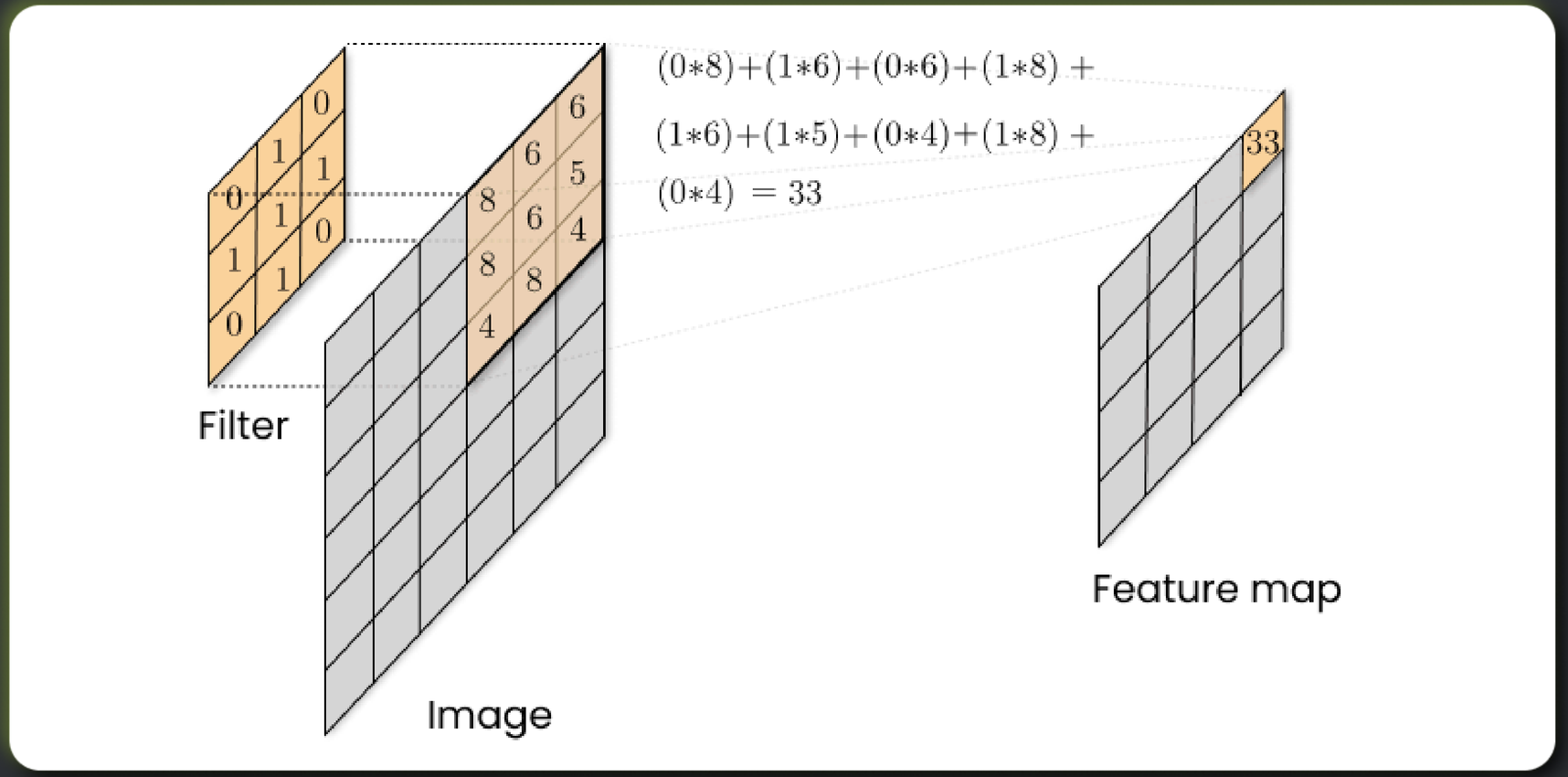
Overview of a machine vision pipeline workflow

Insect centered image

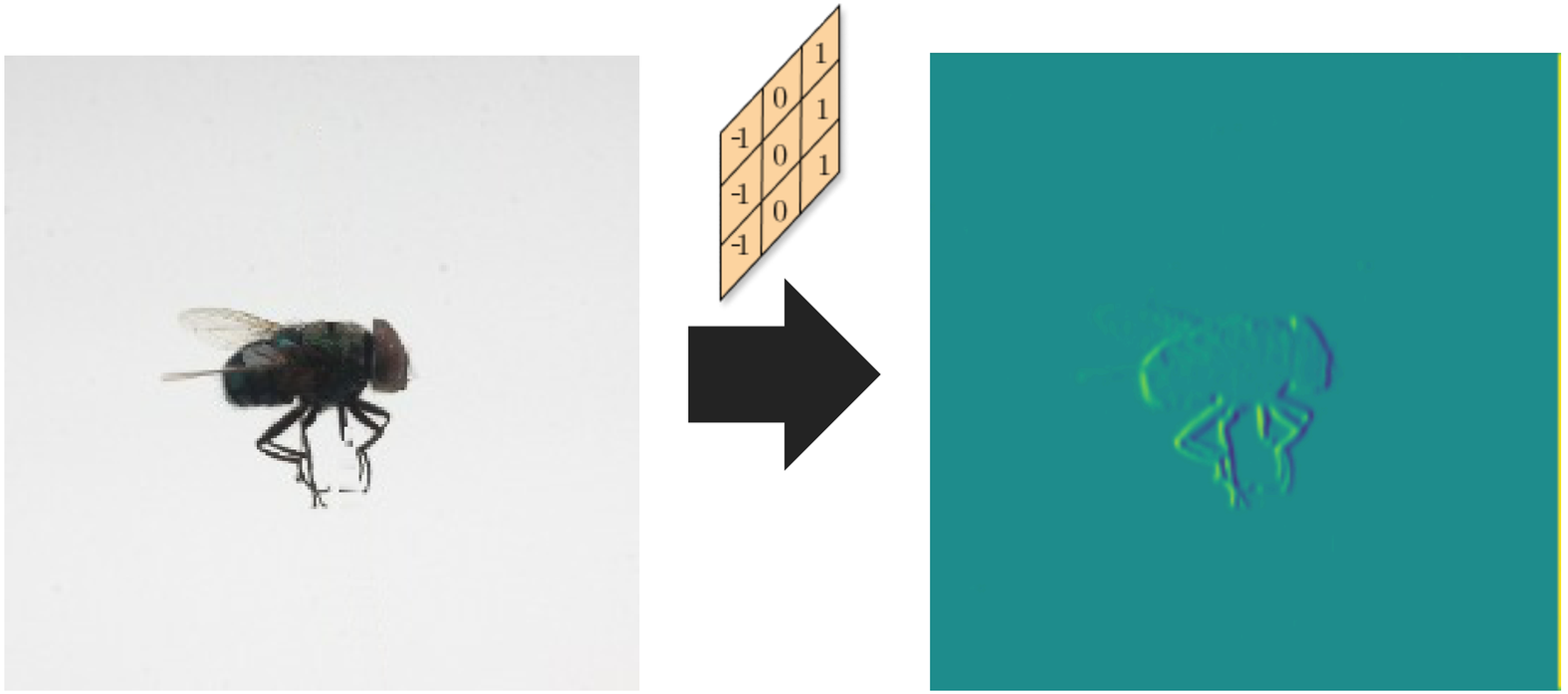
Insect mask



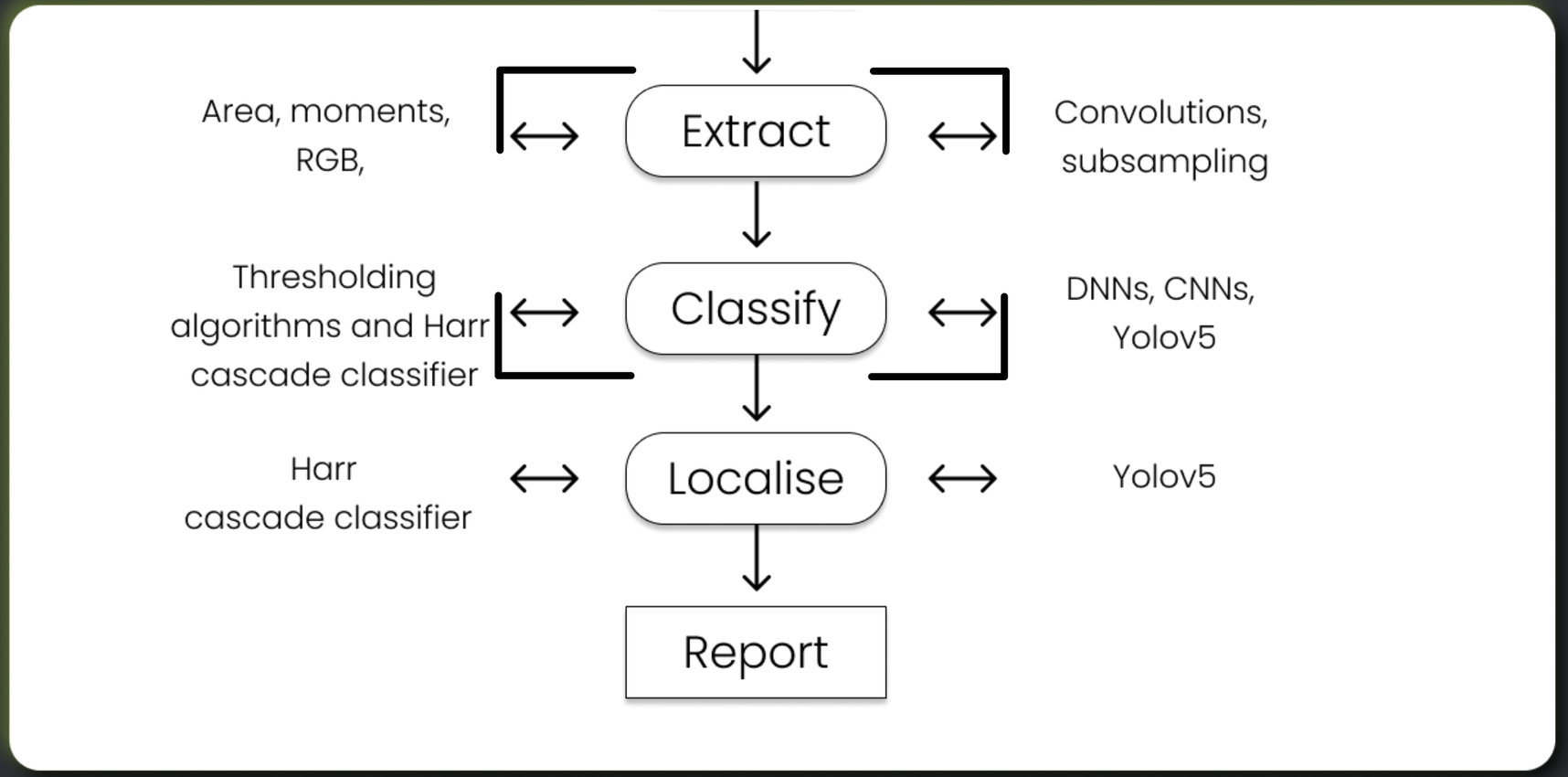
Overview of a machine vision pipeline workflow



Overview of a machine vision pipeline workflow



Overview of a machine vision pipeline workflow

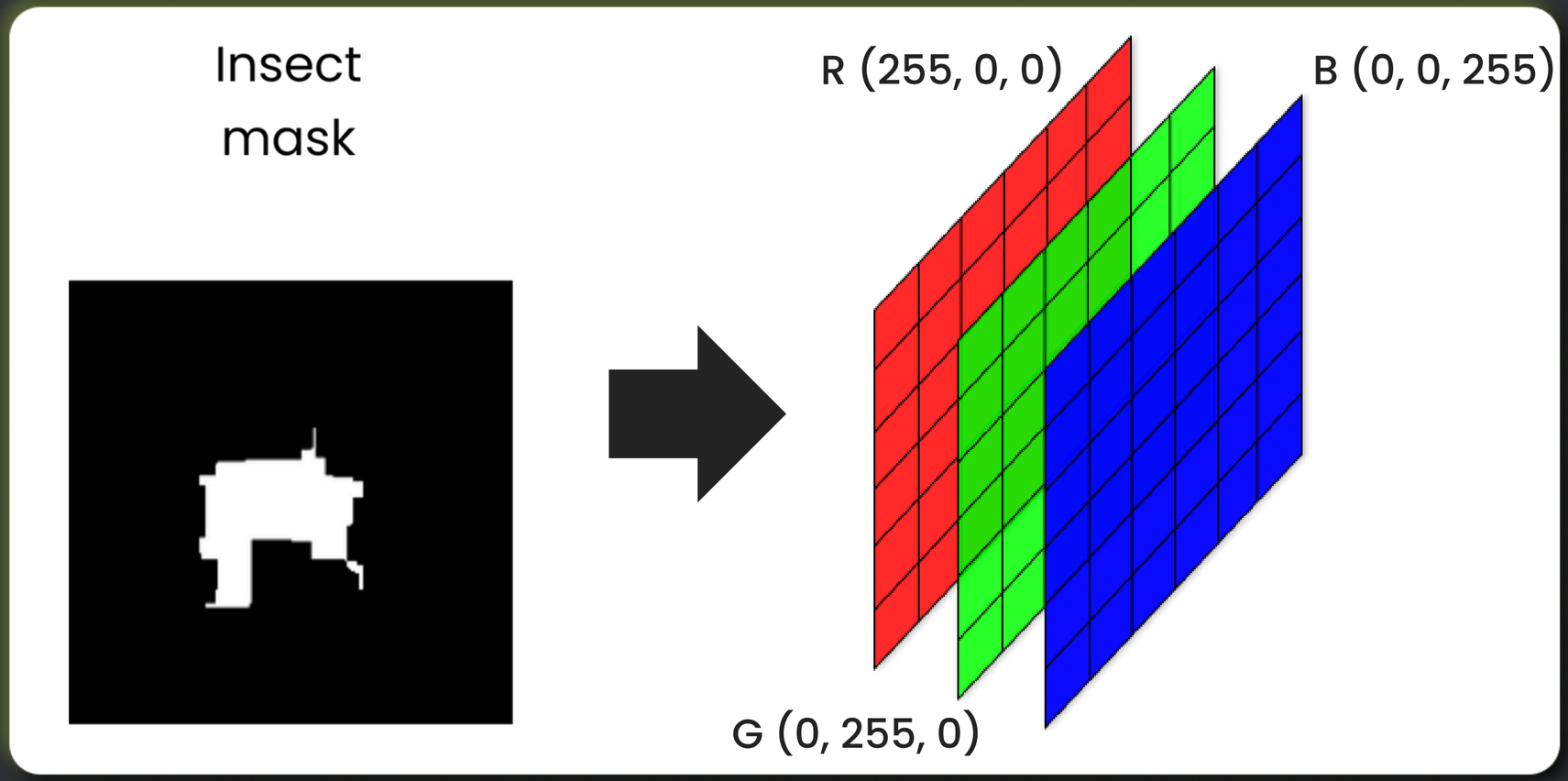


Overview of a machine vision pipeline workflow

Extract: Feature extraction describes each segmented region of the image. Regions of interest have features which include:

- Area;
- Location;
- Eccentricity;
- Centre of gravity (centroid).

Overview of a machine vision pipeline workflow



Overview of a machine vision pipeline workflow

A sample of the feature-based dataset obtained from the medically and forensically important flies dataset (Ong & Ahmad, 2022)

Ratio	Area	$\bar{R}$	$\bar{G}$	$\bar{B}$	$R_{2.5\%}$	$G_{2.5\%}$	$B_{2.5\%}$	$R_{97.5\%}$	$G_{97.5\%}$	$B_{97.5\%}$	Class
0.95	5748	220	219	221	51	56	62	238	238	239	*Lucilia*
1.27	7164	198	194	197	48	36	47	215	215	216	Rhiniinae
0.86	6956	213	211	212	43	37	40	233	233	233	*Chrysomya*
0.41	3500	189	187	188	72	48	63	199	199	199	*Stomorhina*
1.17	7045	208	206	208	52	44	51	226	226	227	*Sarcophaga*

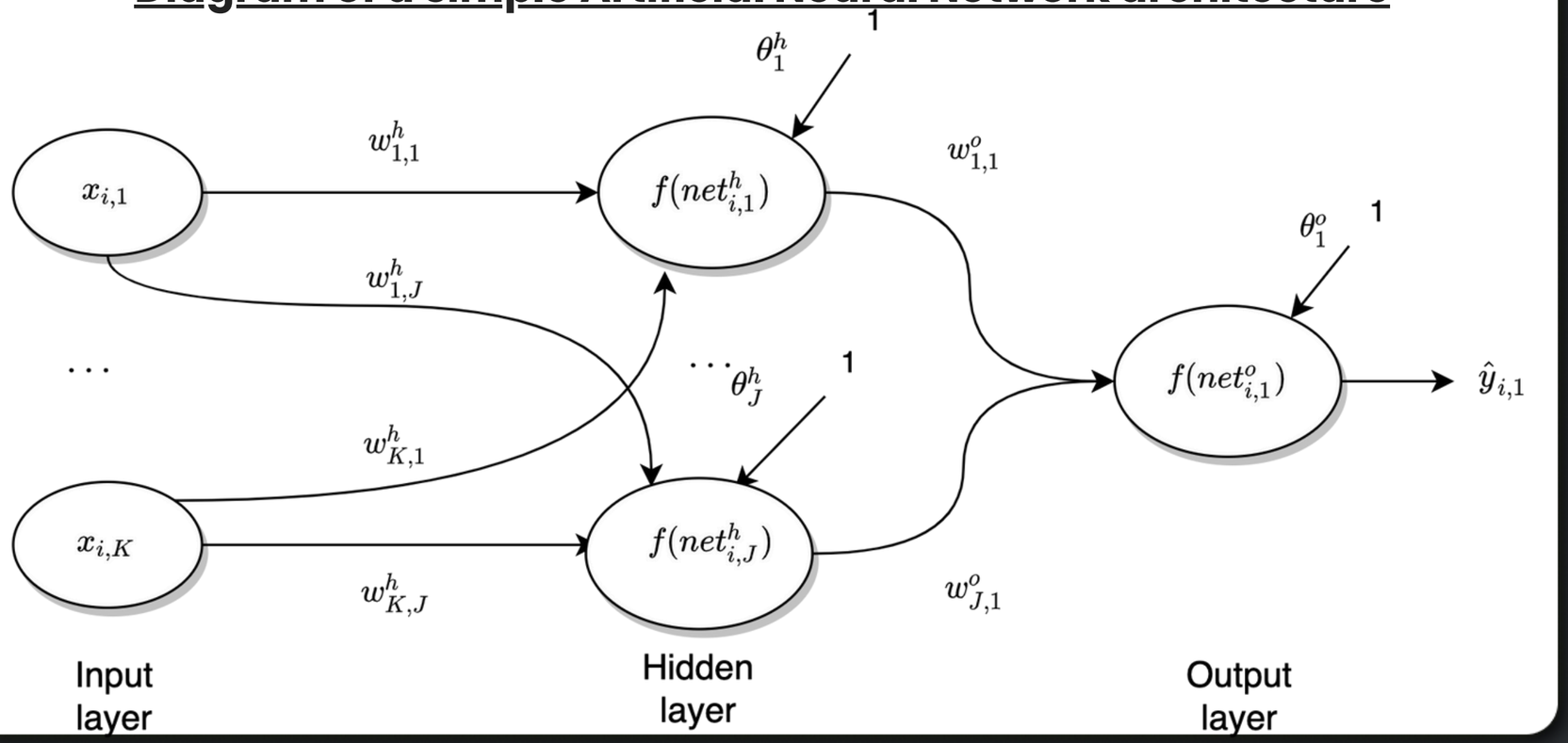
Overview of a machine vision pipeline workflow

Classify: classification reduces the features associated with each segmented region to a single category. Classification provides:

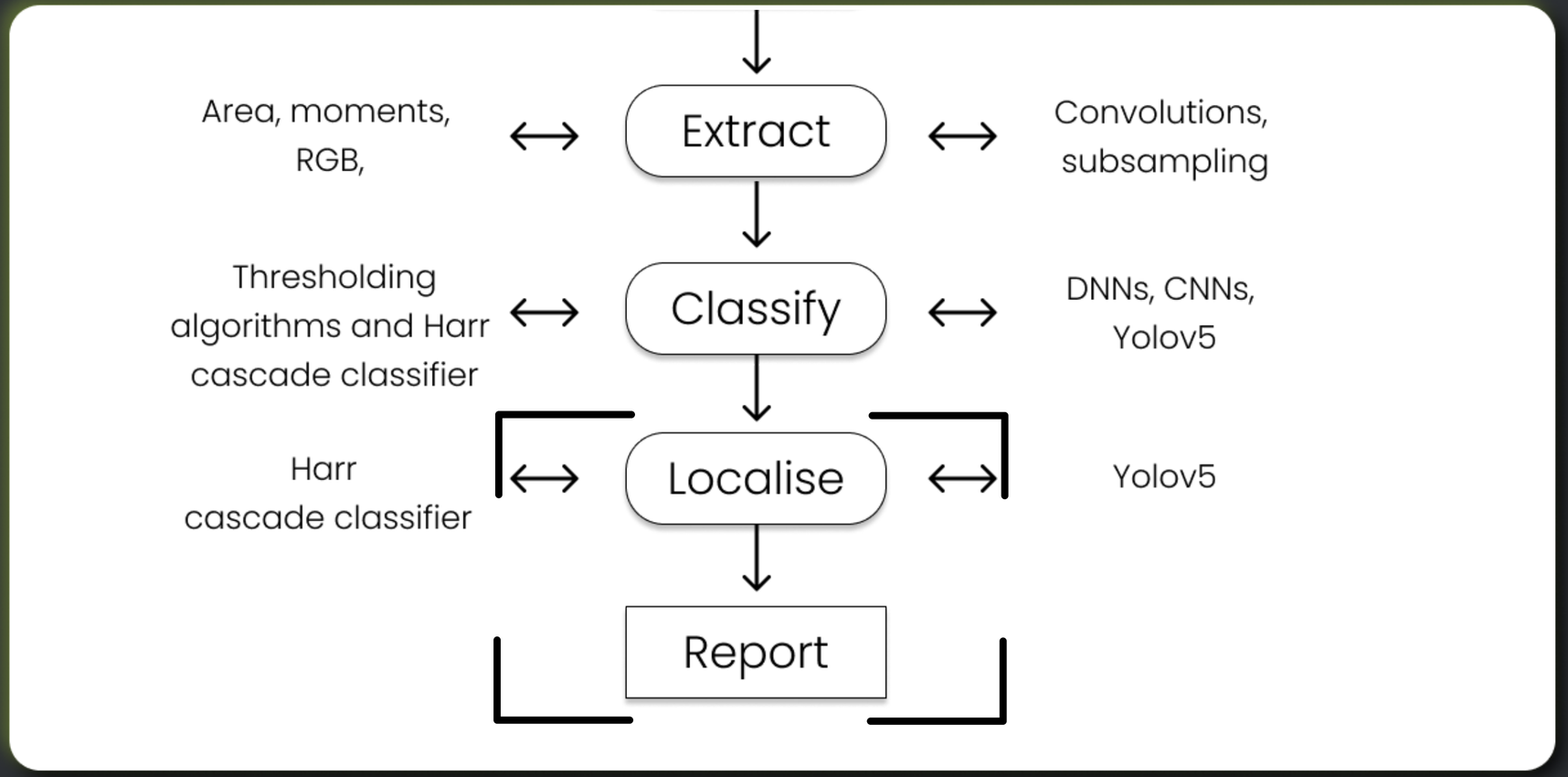
- a semantic label or name (e.g., a species name) associated with the whole image or a region of the image.

Overview of a machine vision pipeline workflow

Diagram of a simple Artificial Neural Network architecture



Overview of a machine vision pipeline workflow



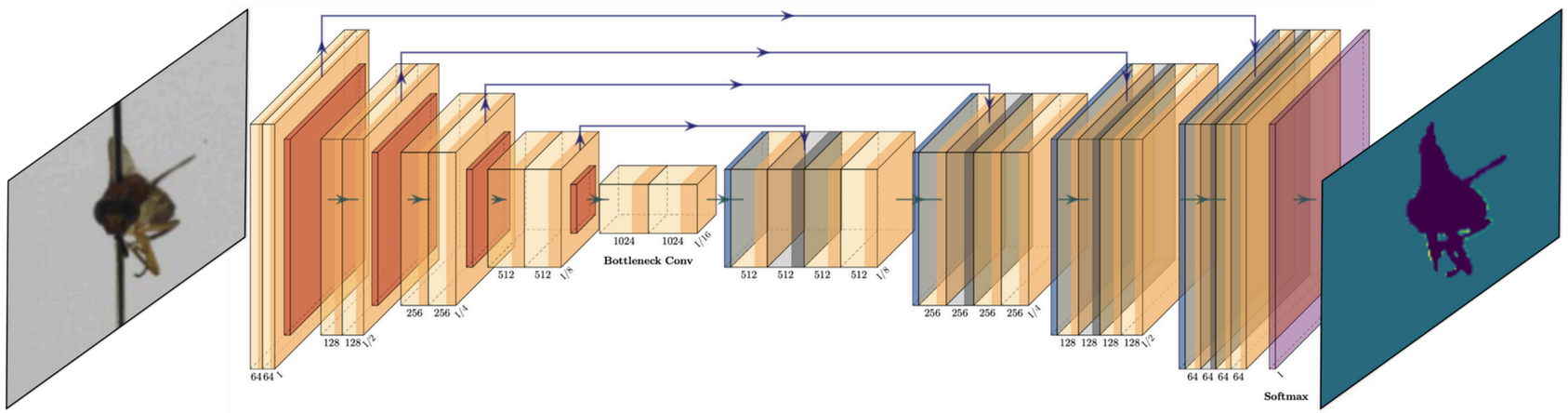
Basic principles of image processing

Localise: returns the position of the classified feature in the image. The additional step of localisation can be computationally intensive. Typically, localisation returns:

- Bounding box (box that surrounds classified features);
- Mask (that can define a classified feature exactly) that intersects each instance of the classified object.

Basic principles of image processing

An example of image segmentation using the from the medically and forensically important flies dataset (Ong & Ahmad, 2022)

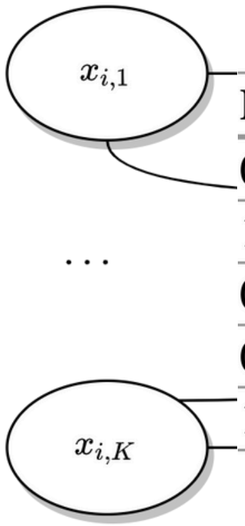


Applications of Deep Learning and Machine Vision

ESALQ

What is a Deep Neural Network?

Simple ANN with $K = 11$ input neurons.

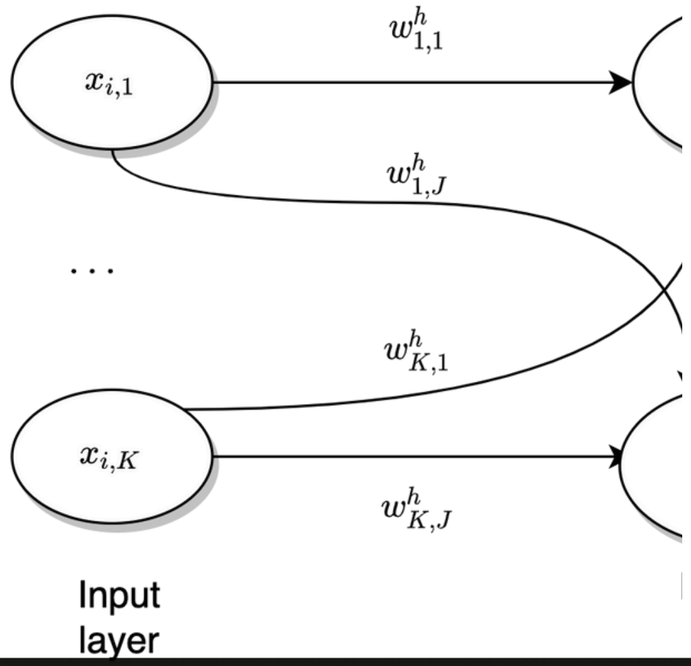


Ratio	Area	$\bar{R}$	$\bar{G}$	$\bar{B}$	$R_{2.5\%}$	$G_{2.5\%}$	$B_{2.5\%}$	$R_{97.5\%}$	$G_{97.5\%}$	$B_{97.5\%}$
0.95	5748	220	219	221	51	56	62	238	238	239
1.27	7164	198	194	197	48	36	47	215	215	216
0.86	6956	213	211	212	43	37	40	233	233	233
0.41	3500	189	187	188	72	48	63	199	199	199
1.17	7045	208	206	208	52	44	51	226	226	227

Input layer

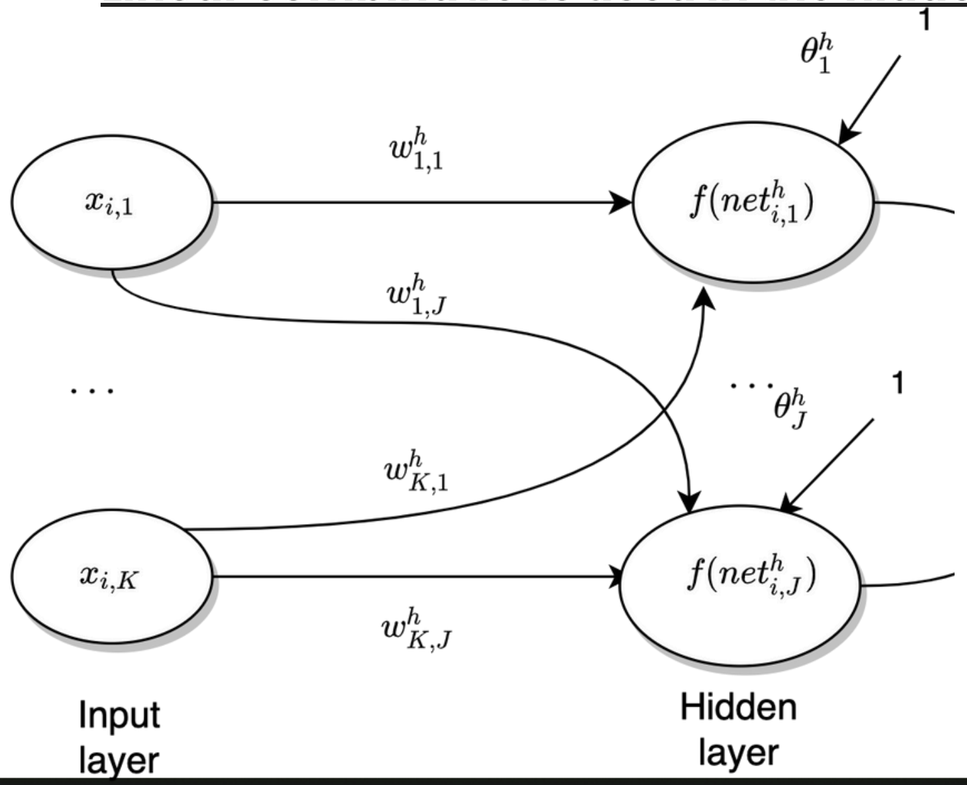
What is a Deep Neural Network?

Weights of a simple ANN connecting the input and hidden layers



What is a Deep Neural Network?

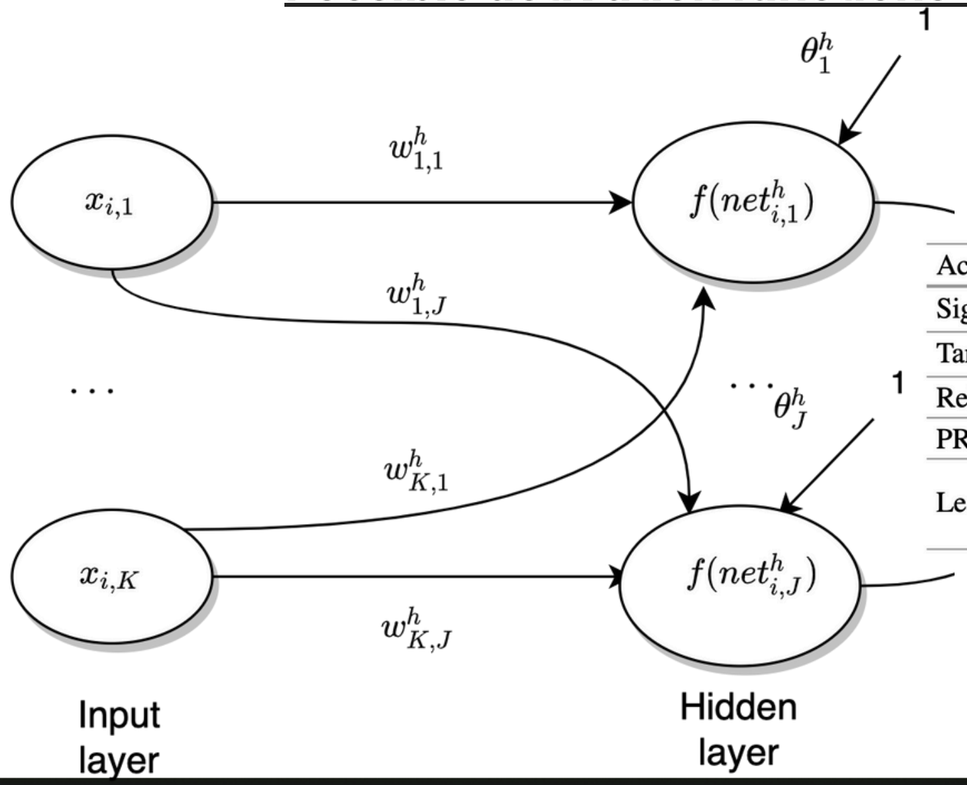
Linear combinations used in the hidden layer of a simple ANN



$$net_{i,j}^h = \sum_k^K w_{k,j}^h x_{i,k} + \theta_j^h$$

What is a Deep Neural Network?

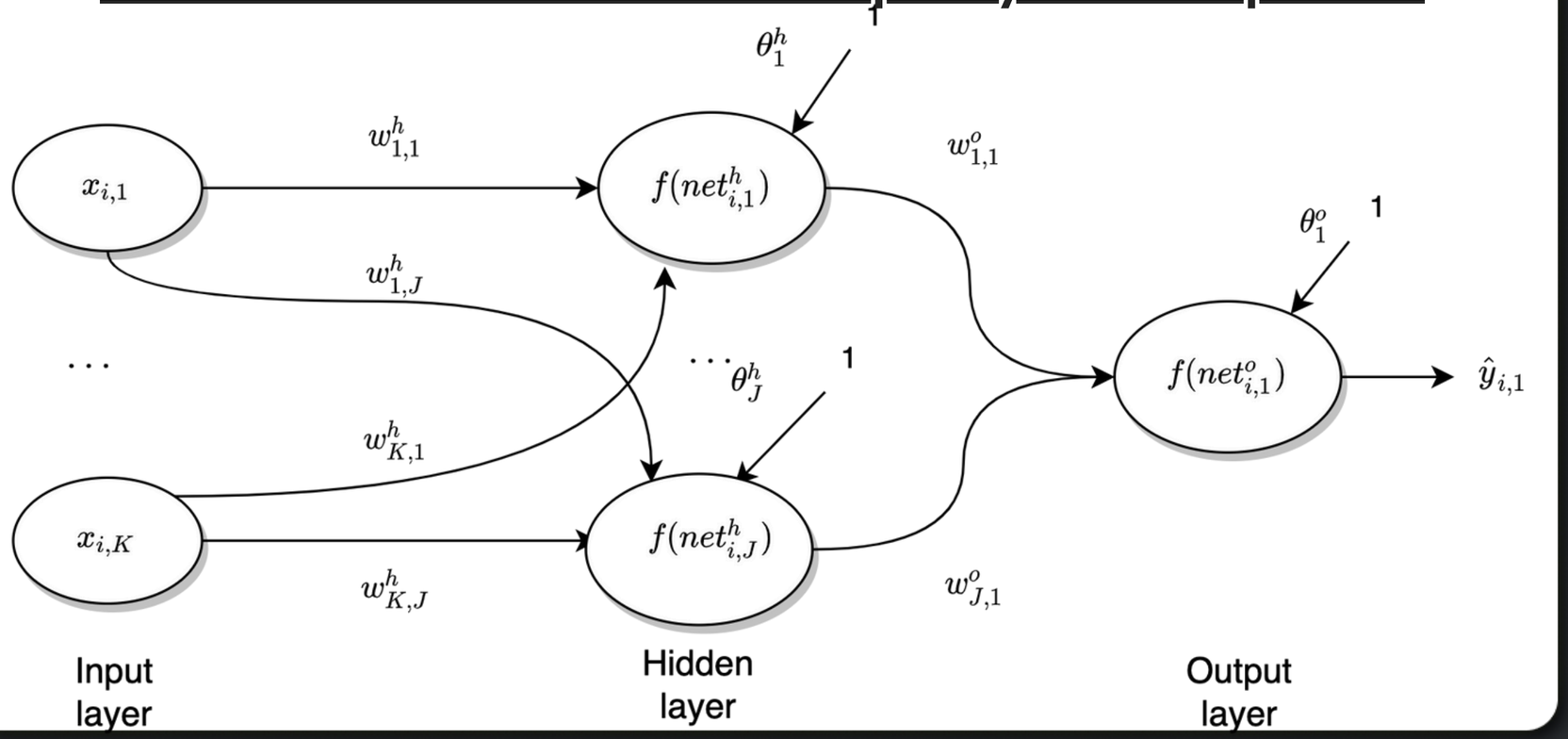
Possible activation functions of a simple ANN



Activation function	Equation
Sigmoid	$f(x) = \frac{1}{1+\exp(-x)}$
Tanh	$f(x) = \frac{1-\exp(-2x)}{1+\exp(-2x)}$
ReLU	$f(x) = \max(0, x)$
PReLU	$f(x) = \max(0, x) + \alpha \min(0, x)$
Leaky ReLU	$f(x) = \begin{cases} \max(0, x), & \text{if } x \geq 0 \\ \alpha \max(0, x), & \text{otherwise} \end{cases}$

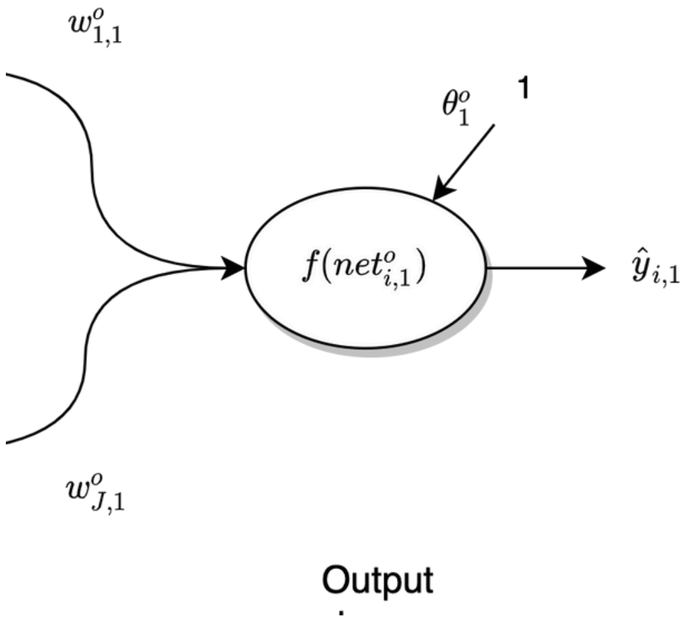
What is a Deep Neural Network?

Linear combinations used in the output layer of a simple ANN



What is a Deep Neural Network?

Linear combinations used in the output layer of a simple ANN



$$net^o_{i,1} = \sum_j^J w^o_{j,1} f(net^h_{i,j}) + \theta^o_1$$

What is a Deep Neural Network?

Learning phase of a simple ANN

$$E_i^2 = \left(y_{i,1} - f(net_{i,1}^o)\right)^2 = \left(y_{i,1} - f\left(\sum_j^J f(net_{i,j}^h) w_{k,j}^o + \theta_1^o\right)\right)^2$$

What is a Deep Neural Network?

Learning weight parameters of a simple ANN

$$w_{k,j}^{h}(t+1) = w_{k,j}^{h}(t) - \eta \frac{\partial E_i^2}{\partial w_{k,j}^{h}}$$

$$w_{k,j}^{o}(t+1) = w_{k,j}^{o}(t) - \eta \frac{\partial E_i^2}{\partial w_{k,j}^{o}}$$

What is a Deep Neural Network?

Learning bias parameters of a simple ANN

$$\theta_j^h(t+1) = \theta_j^h - \eta \frac{\partial E_i^2}{\partial \theta_j^h}$$

$$\theta_j^h(t+1) = \theta_j^o - \eta \frac{\partial E_i^2}{\partial \theta_j^o}$$

Applications of Deep Learning and Machine Vision

ESALQ

Overview

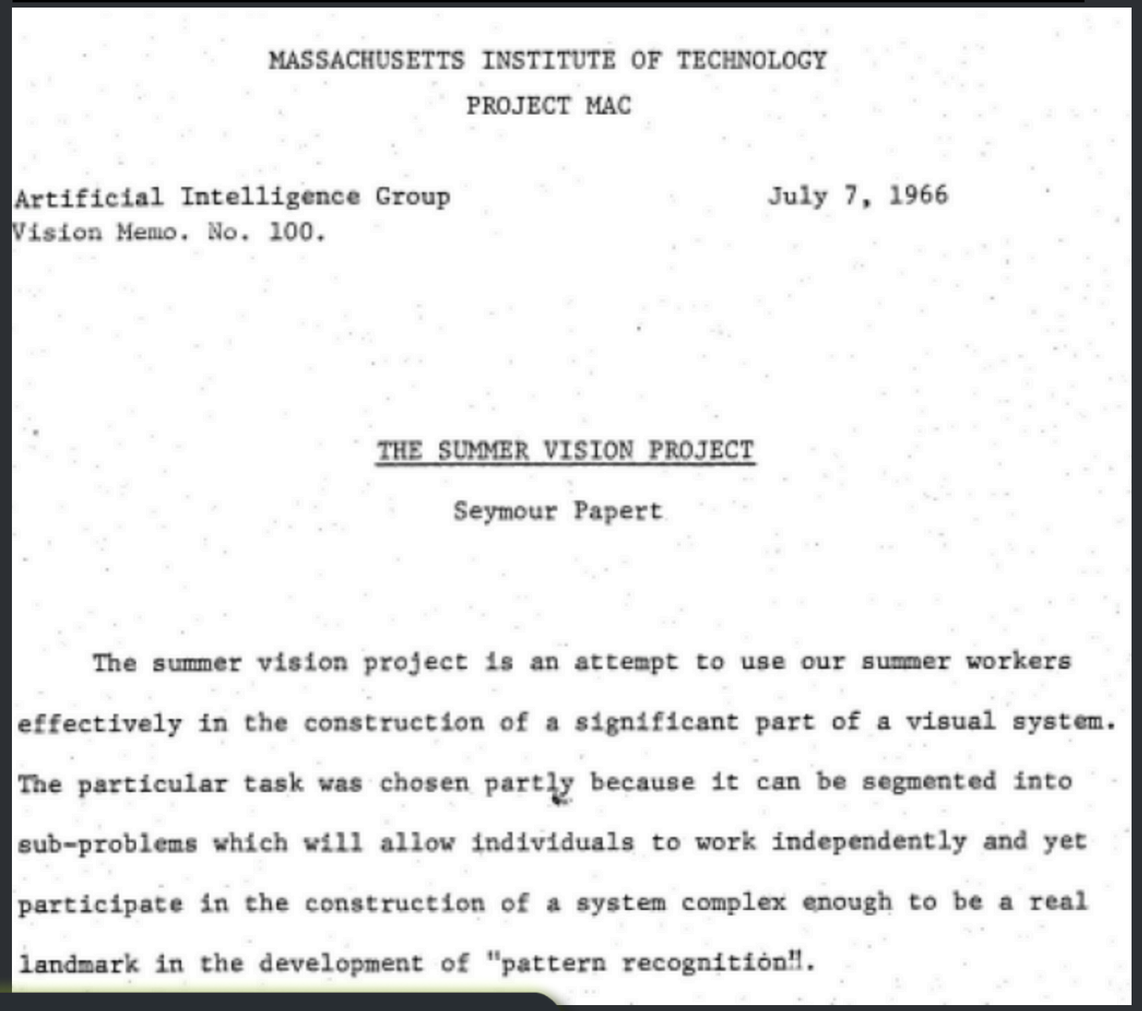
Day 1

- An introduction to Machine Vision:
  - Overview of a machine vision pipeline workflow;
  - An introduction to OpenCV in Python;
- An introduction to Artificial Neural Networks:
  - What is a Deep Neural Network?
  - DNN using TensorFlow

Day 2

- An introduction to Convolutional Neural Networks;
  - What are the basic operations of a CNN?
  - How do we train a CNN?
- An introduction to Transfer Learning;
  - How can we take advantage of transfer learning?

An introduction to Convolutional Neural Networks



MASSACHUSETTS INSTITUTE OF TECHNOLOGY

PROJECT MAC

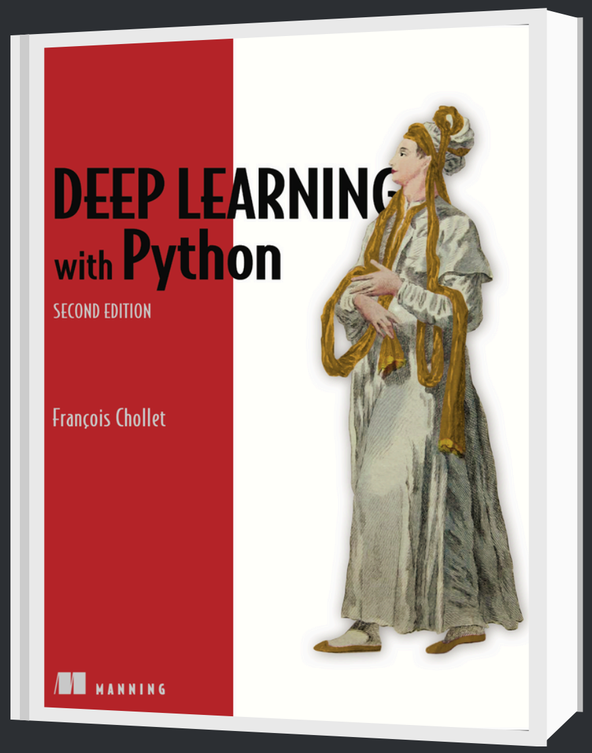
Artificial Intelligence Group — July 7, 1966
Vision Memo. No. 100.

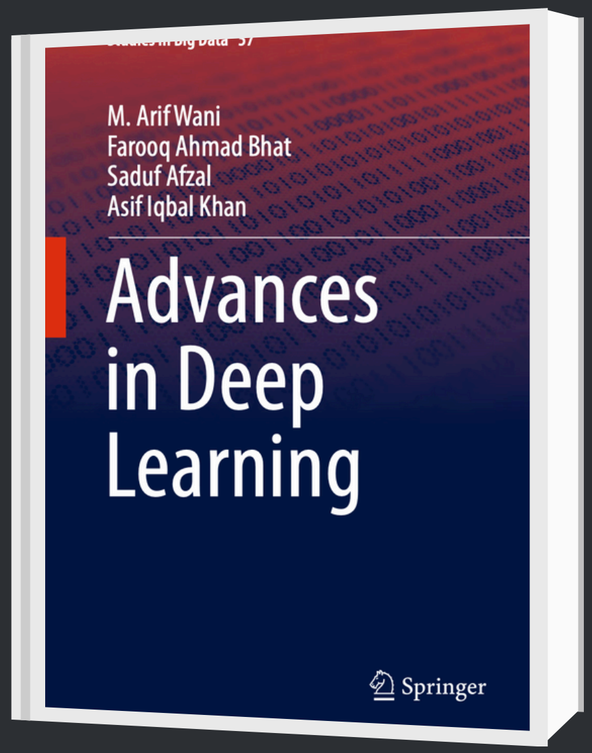
THE SUMMER VISION PROJECT

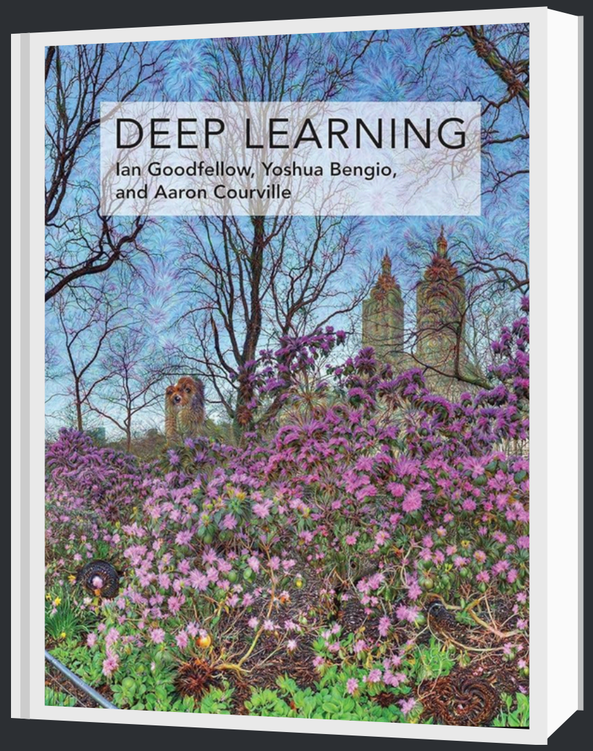
Seymour Papert

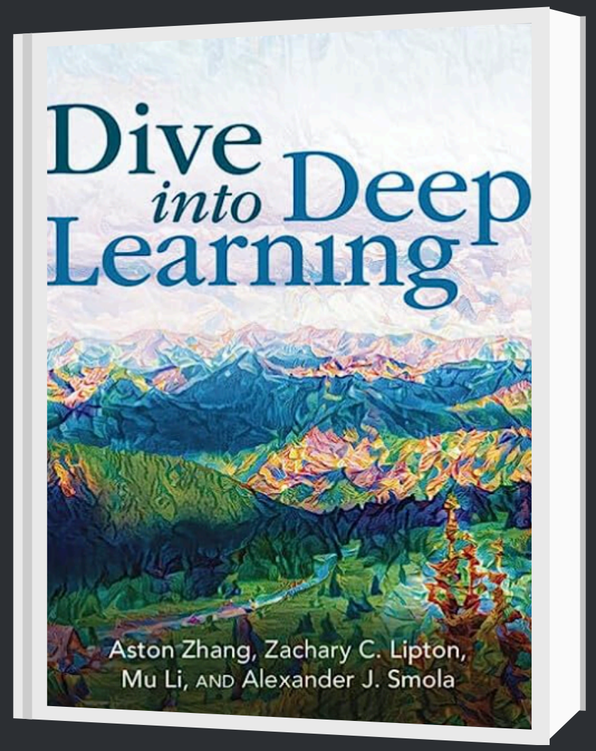
The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".

An introduction to Machine Vision









Applications of Deep Learning and Machine Vision

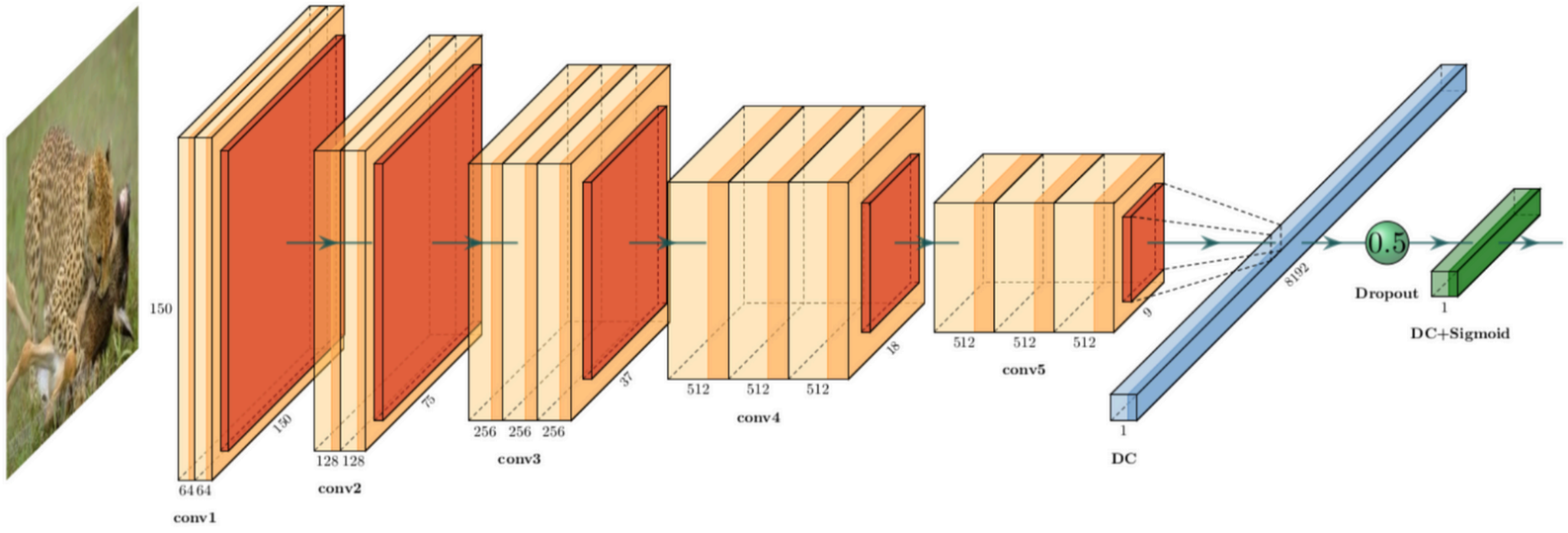
ESALQ

Applications of Deep Learning and Machine Vision

ESALQ

An introduction to Convolutional Neural Networks

Detecting predation interaction using pretrained CNNs



An introduction to Convolutional Neural Networks

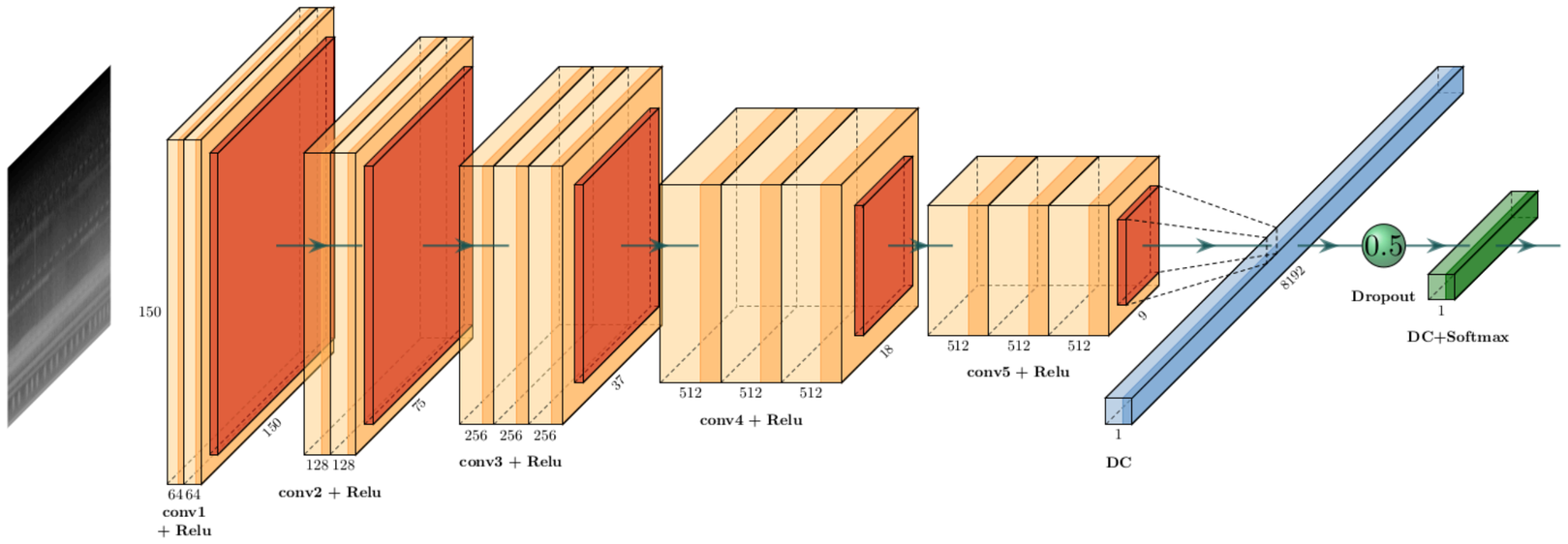
Detecting predation interaction using pretrained CNNs

Applications of Deep Learning and Machine Vision

ESALQ

An introduction to Convolutional Neural Networks

A machine vision system for avian song classification with CNN's



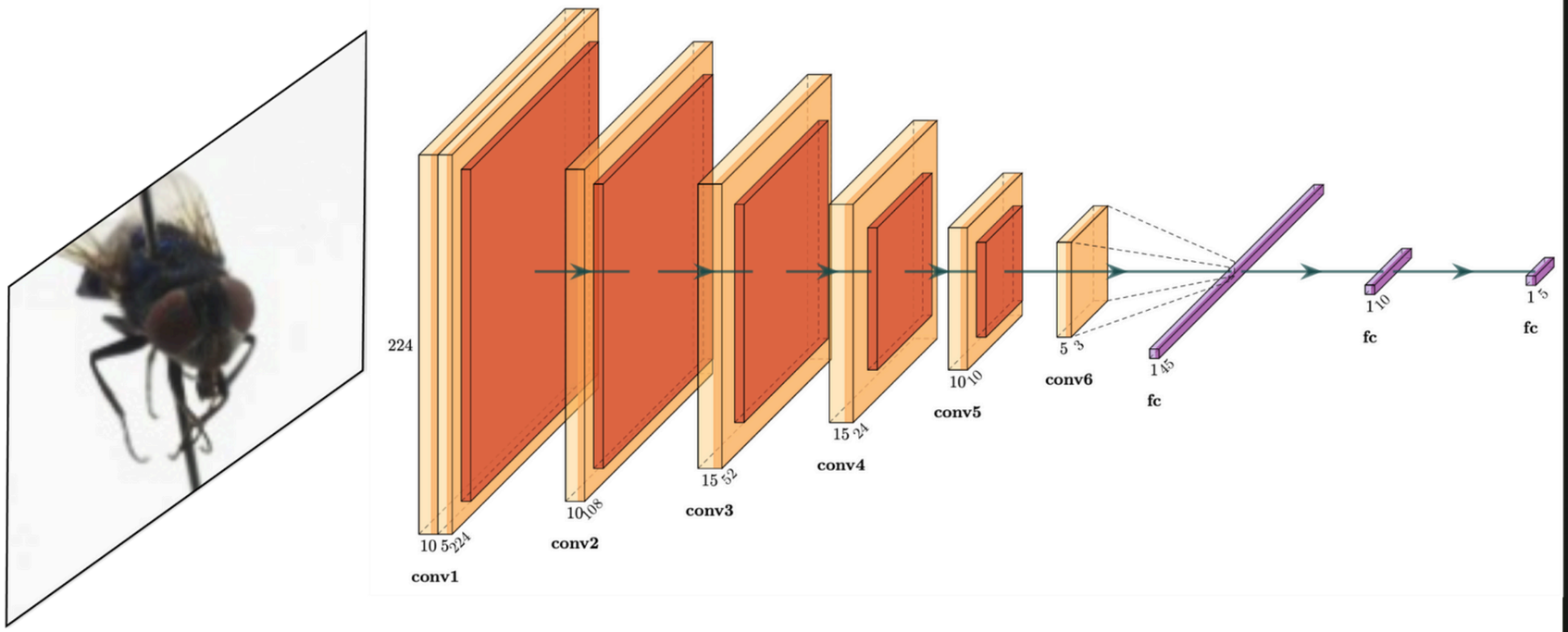
ESALQ

Applications of Deep Learning and Machine Vision

ESALQ

An introduction to Convolutional Neural Networks

Machine Vision Applied to Entomology

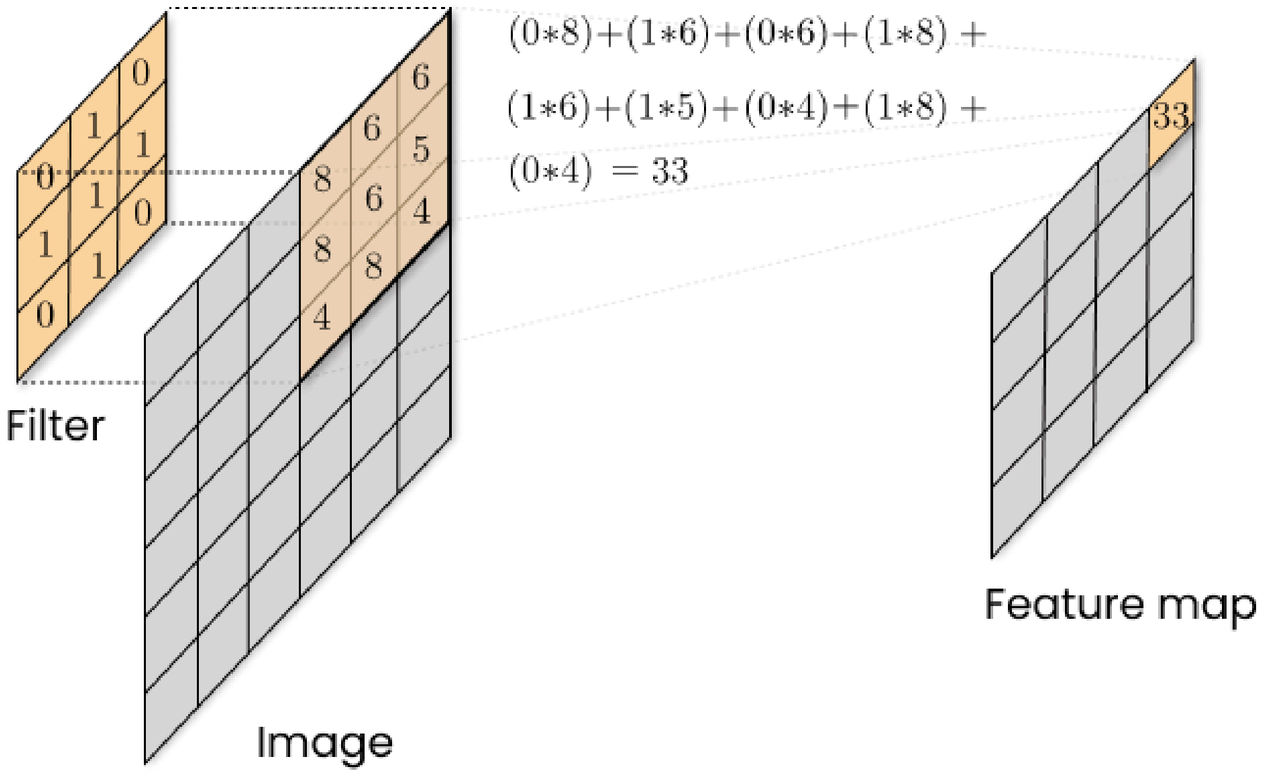


ESALQ

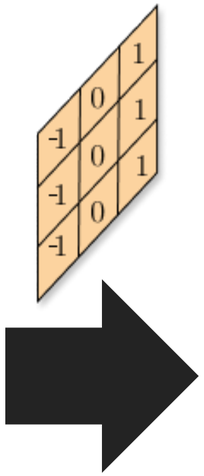
Applications of Deep Learning and Machine Vision

ESALQ

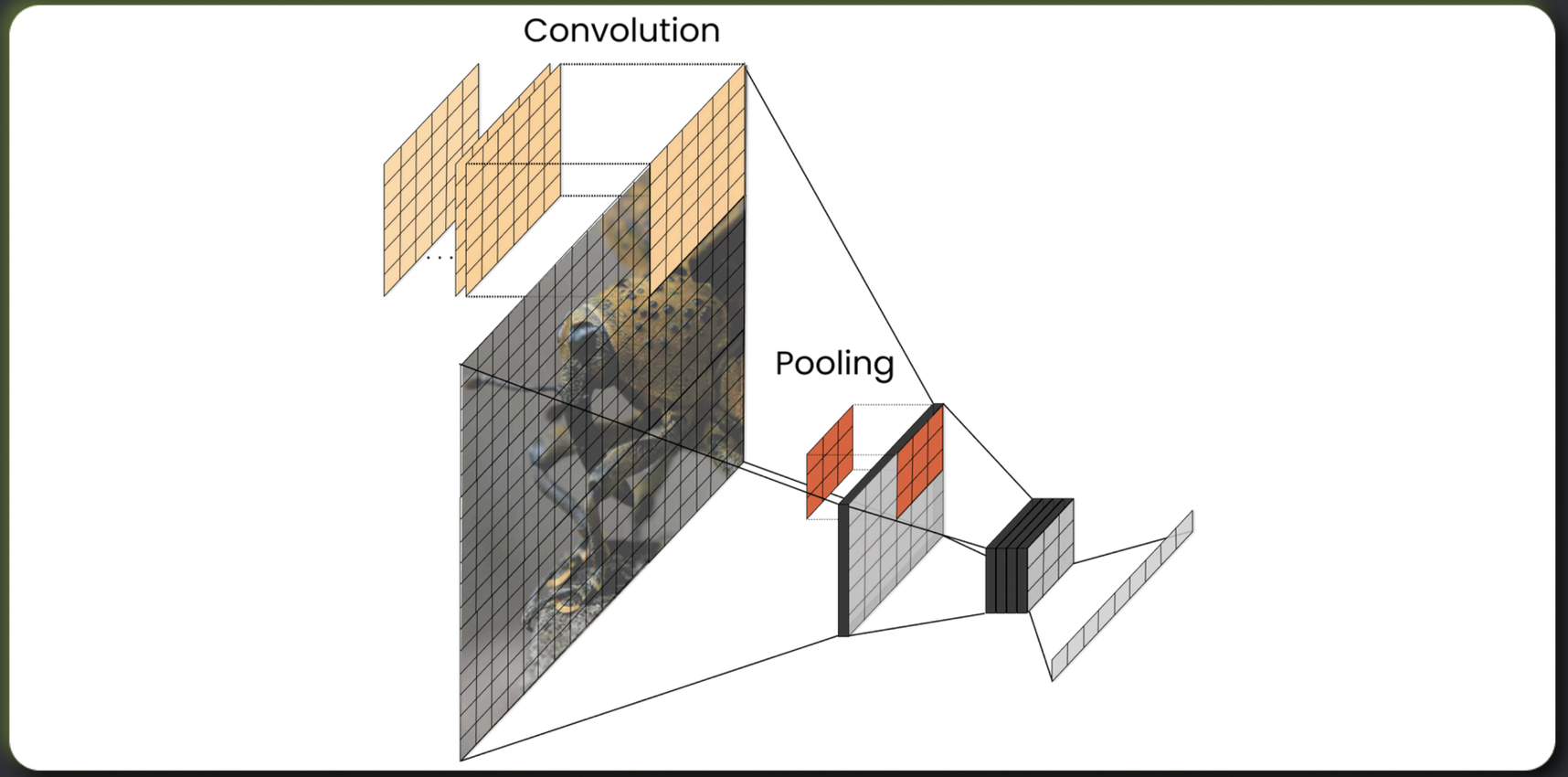
What are the basic operations of a CNN?



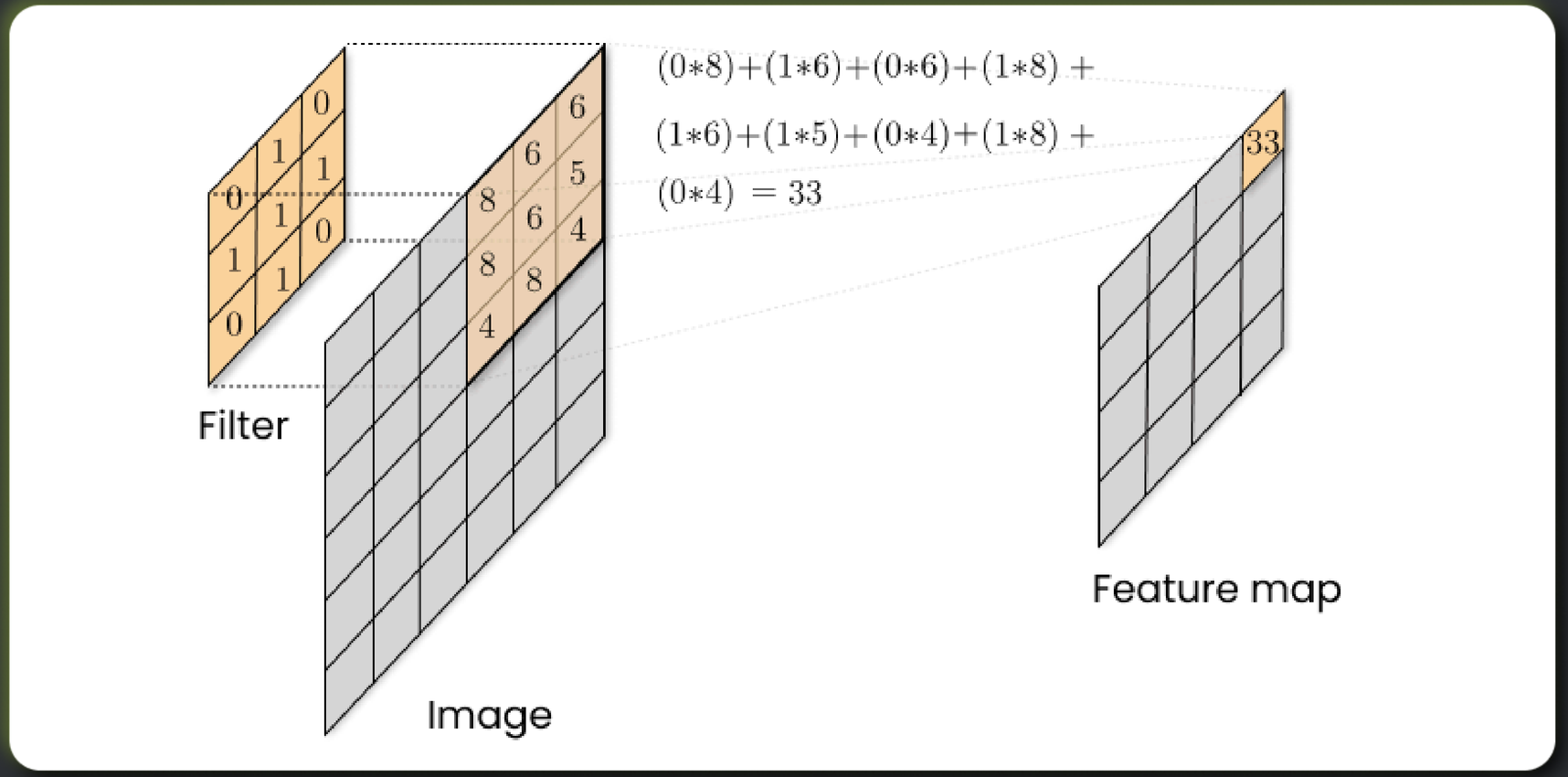
What are the basic operations of a CNN?



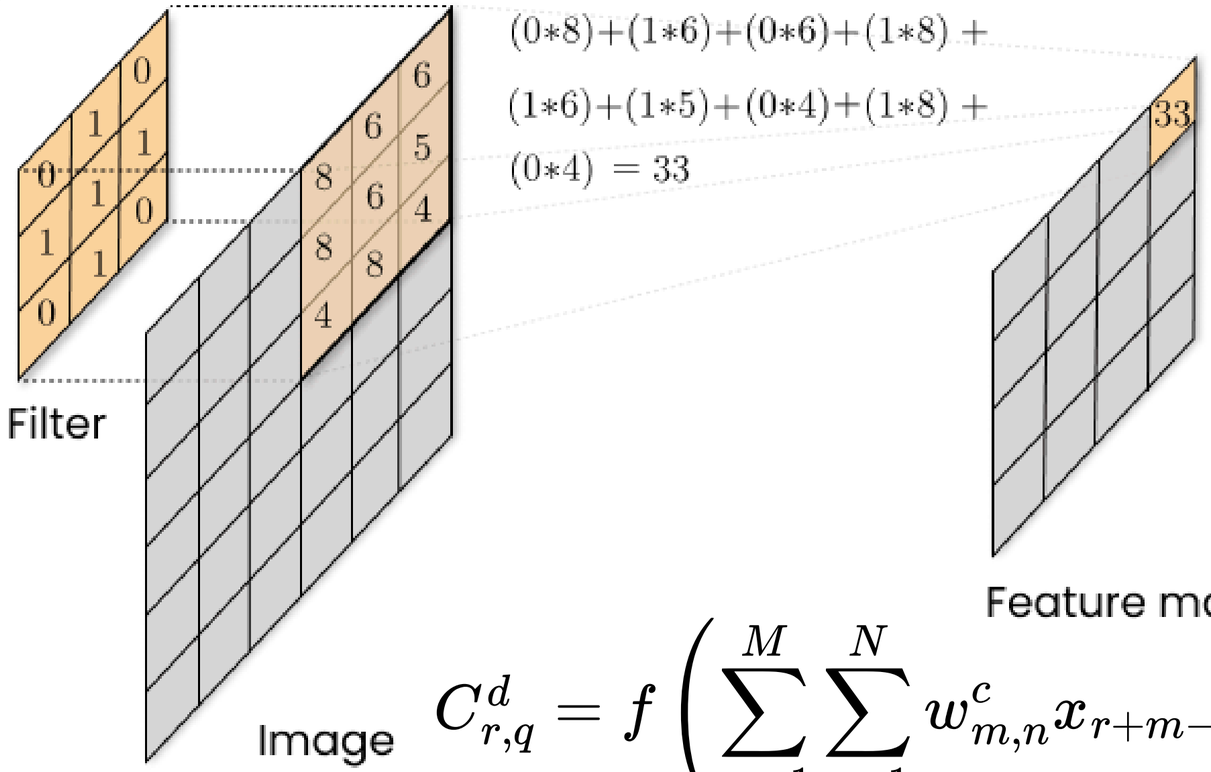
What are the basic operations of a CNN?



What are the basic operations of a CNN?

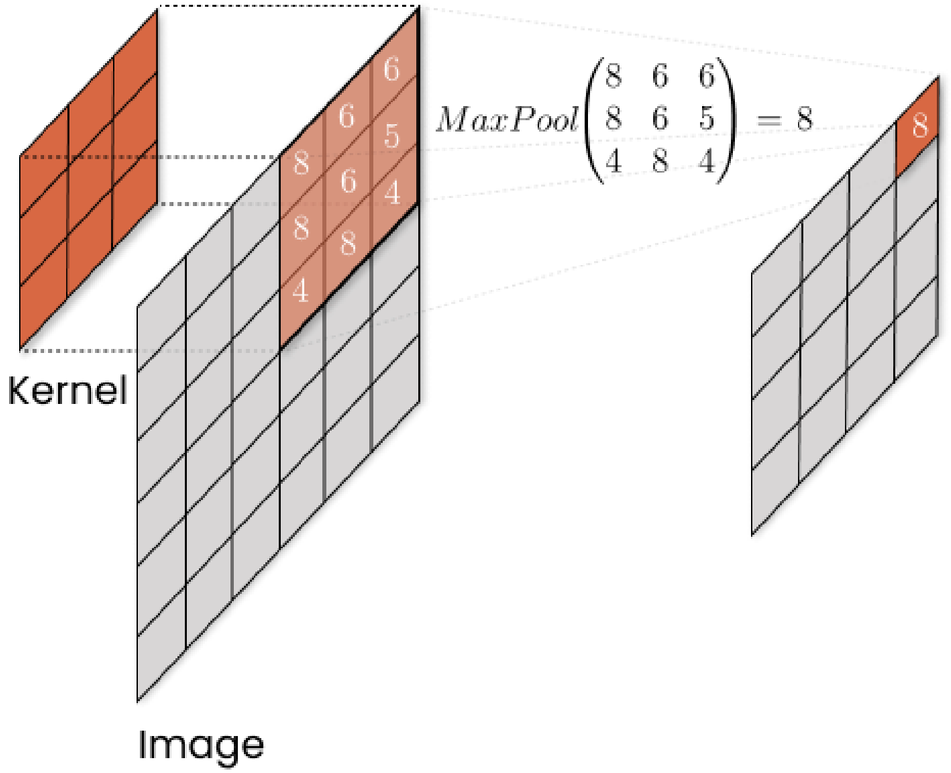


What are the basic operations of a CNN?

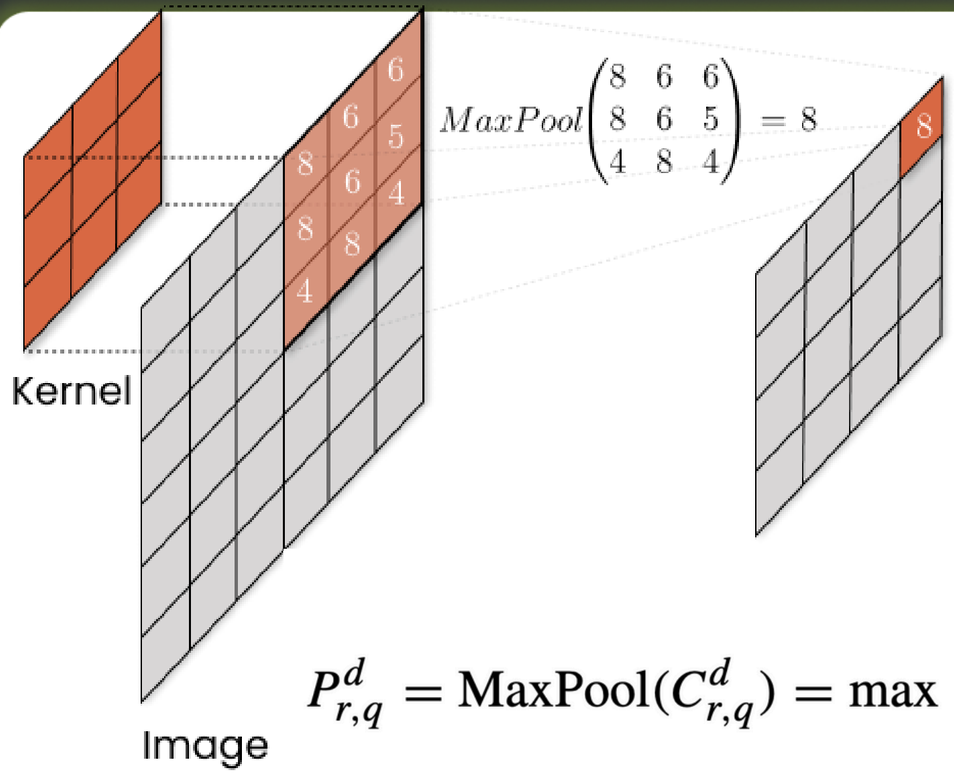


$$C^d_{r,q} = f\left(\sum_{m=1}^{M}\sum_{n=1}^{N} w^c_{m,n} x_{r+m-1,q+n-1} + \theta^c\right)$$

What are the basic operations of a CNN?



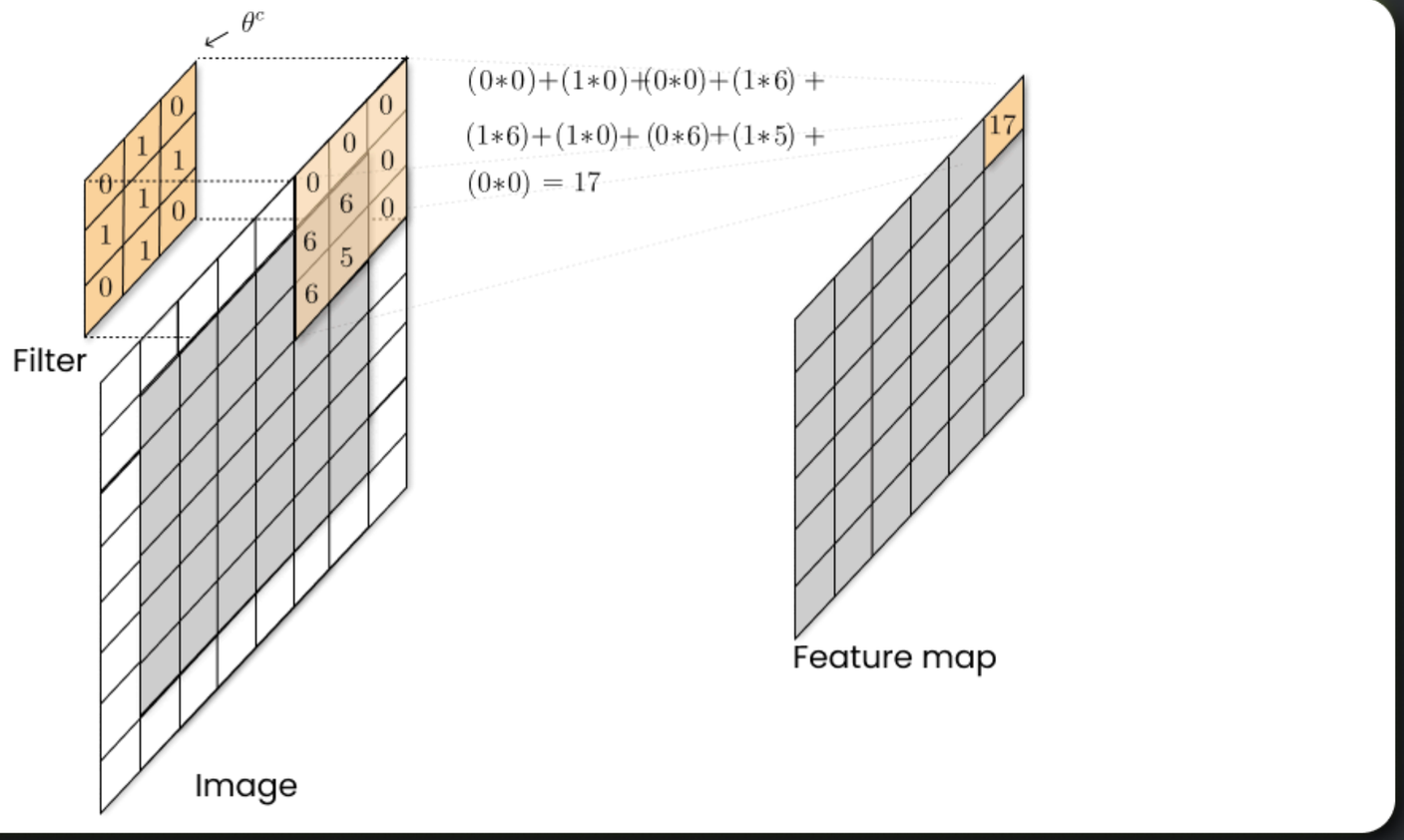
What are the basic operations of a CNN?



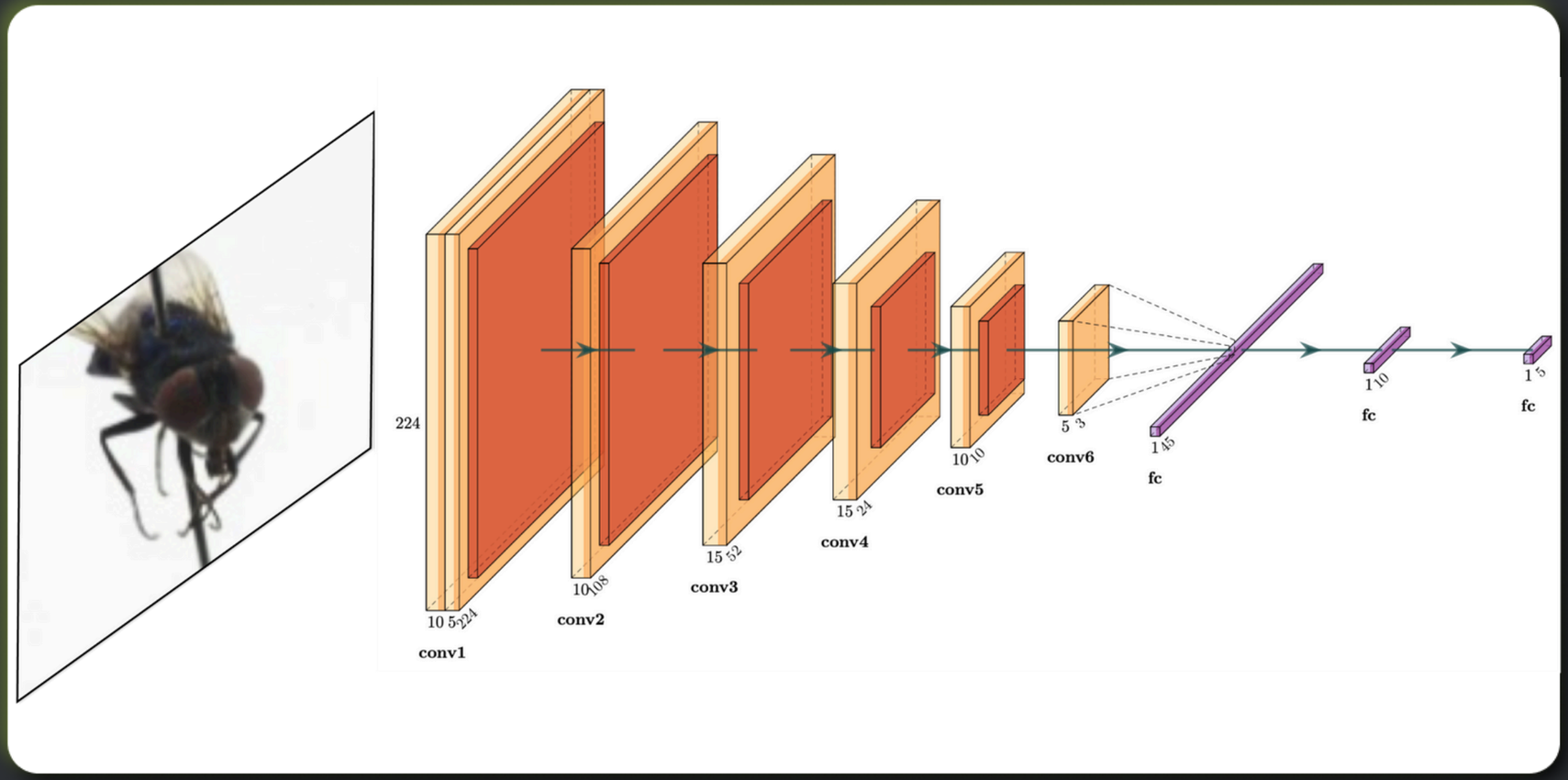
$$MaxPool\begin{pmatrix} 8 & 6 & 6 \\ 8 & 6 & 5 \\ 4 & 8 & 4 \end{pmatrix} = 8$$

$$P^d_{r,q} = \text{MaxPool}(C^d_{r,q}) = \max \begin{pmatrix} C^d_{r,q} & C^d_{r+1,q} & C^d_{r+2,q} \\ C^d_{r+1,q} & C^d_{r+1,q+1} & C^d_{r+2,q} \\ C^d_{r+2,q} & C^d_{r+2,q+1} & C^d_{r+2,q+2}, \end{pmatrix}$$

What are the basic operations of a CNN?



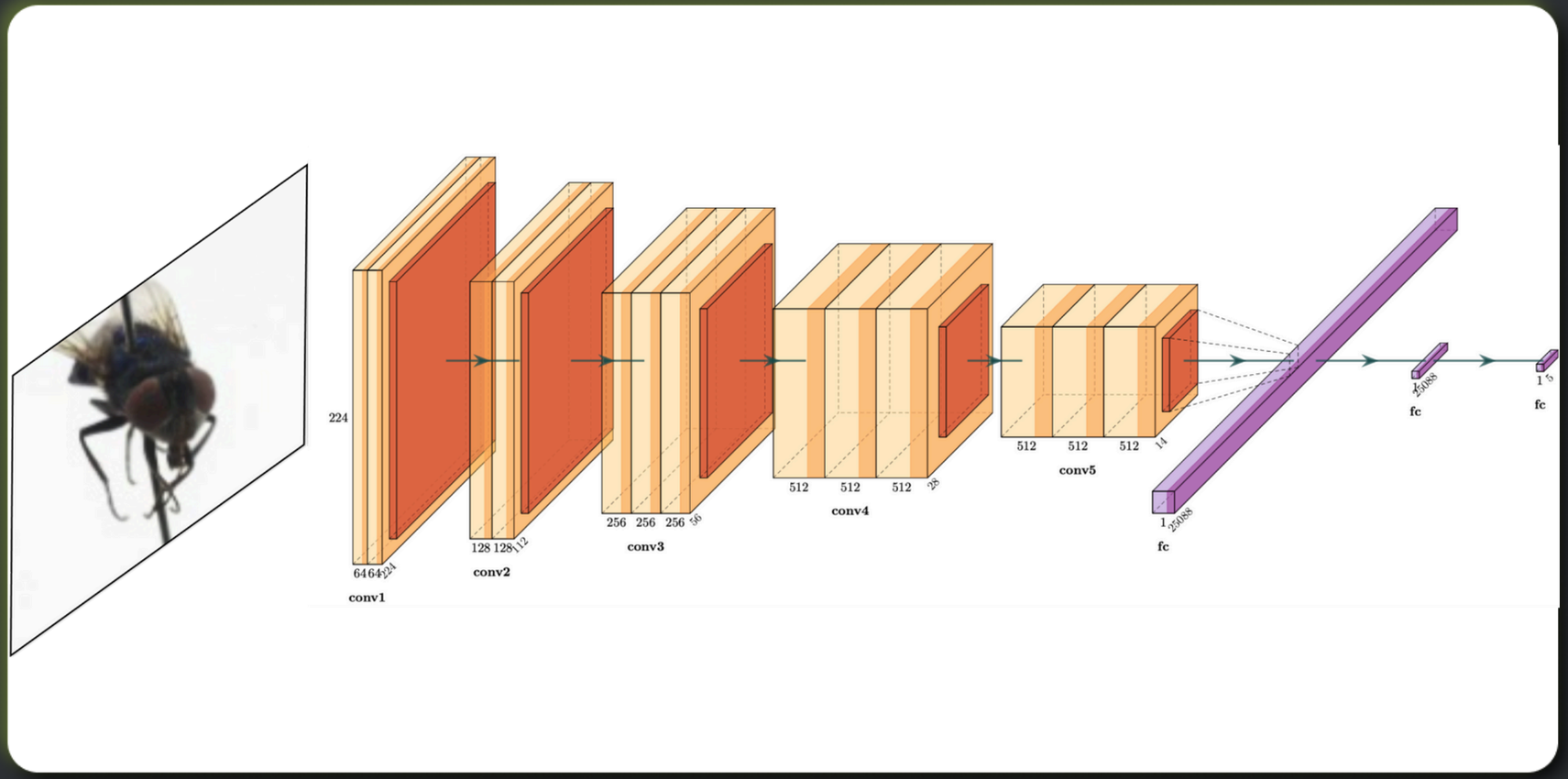
What are the basic operations of a CNN?



Applications of Deep Learning and Machine Vision

ESALQ

How can we take advantage of transfer learning?



How can we take advantage of transfer learning?

Applications of Deep Learning and Machine Vision

ESALQ