Image Processing and its Applications

MID-TERM EVAL

Team: Entropy

BY
Dhruv Bajaj
Rutul Bhanushali
Nonit Gupta
Bipin Kumar Jiaswal

INTRODUCTION.

What is Image Processing?

- Definition:
  - Image processing is the technique of manipulating and analyzing digital images using algorithms to enhance or extract meaningful information.
- Key Steps:
  - Input: Raw image data
  - Processing: Enhancement, filtering, transformation
  - Output: Improved image or extracted data
- Goal: Improve visual quality or prepare images for interpretation (by humans or machines).

Applications of Image Processing

- Everyday Examples:
  - Photo editing apps (e.g., filters on Instagram)
  - Facial recognition (e.g., unlocking smartphones)
- Industry Use Cases:
  - Medical imaging (e.g., MRI analysis)
  - Autonomous vehicles (e.g., object detection)
  - Surveillance (e.g., motion tracking)
- Scientific Research:
  - Astronomy (e.g., analyzing telescope images)
  - Environmental monitoring (e.g., satellite imagery)

Week 1 Summary.

NumPy Arrays for Image Data

- Images as Arrays
  - Representation:
    - Grayscale images: 2D arrays (height, width)
    - Color images: 3D arrays (height, width, channels) e.g., RGB

Key Properties

- Shape: Dimensions of the array
- Data Type: Typically uint8 for image data, ensuring proper pixel value handling

In a practical, hands-on context, you can dive into creating arrays that replicate the way an image might be initialized, perhaps by setting up a grid filled entirely with zeros to represent a blank slate, or using ones for a uniform starting point, or even generating random values to simulate noise or texture.

The Week 1 material on NumPy—covered through notebooks, online resources, and assignments—provides crucial skills for image processing with OpenCV. We studied how images are represented as arrays, manipulated (by accessing, modifying, and reshaping), and advanced operations such as normalization, channel manipulation, and filtering

WEEK 2 SUMMARY

What is thresholding?

Thresholding is a technique in which we move our pixel values to the threshold value. The thresholding basically takes each pixel value and compared it with the threshold value. If the pixel value is smaller than the threshold, it is set to 0, otherwise, it is set to a maximum value i.e. 255. It is useful for object detection and segmentation

Types of Thresholding:

Binary Thresholding: Binary Thresholding is a simple image segmentation technique that converts a grayscale image into a black and white (binary) image.

Adaptive mean Thresholding: Adaptive Mean Thresholding is a type of adaptive thresholding where the threshold value for each pixel is calculated as the mean of the neighboring pixel values in a block around it.

Adaptive Gaussian Thresholding: It is a binarization technique where the threshold is calculated for each pixel based on its neighborhood, using a Gaussian weighted sum.

Otsu’s Thresholding: It is an automatic global thresholding method used to convert a grayscale image to binary without manually choosing a threshold.

What is Morphological Transformation?

Morphological transformations are some simple operations based on the image shape. It is normally performed on binary images. It needs two inputs, one is our original image second one is called structuring element or kernel which decides the nature of operation.

Basic operations in Morphological Transformation:

Erosion: It Shrinks the boundaries of foreground object (removes noise).

Dilation: It expands the boundaries of foreground object(fill gaps).

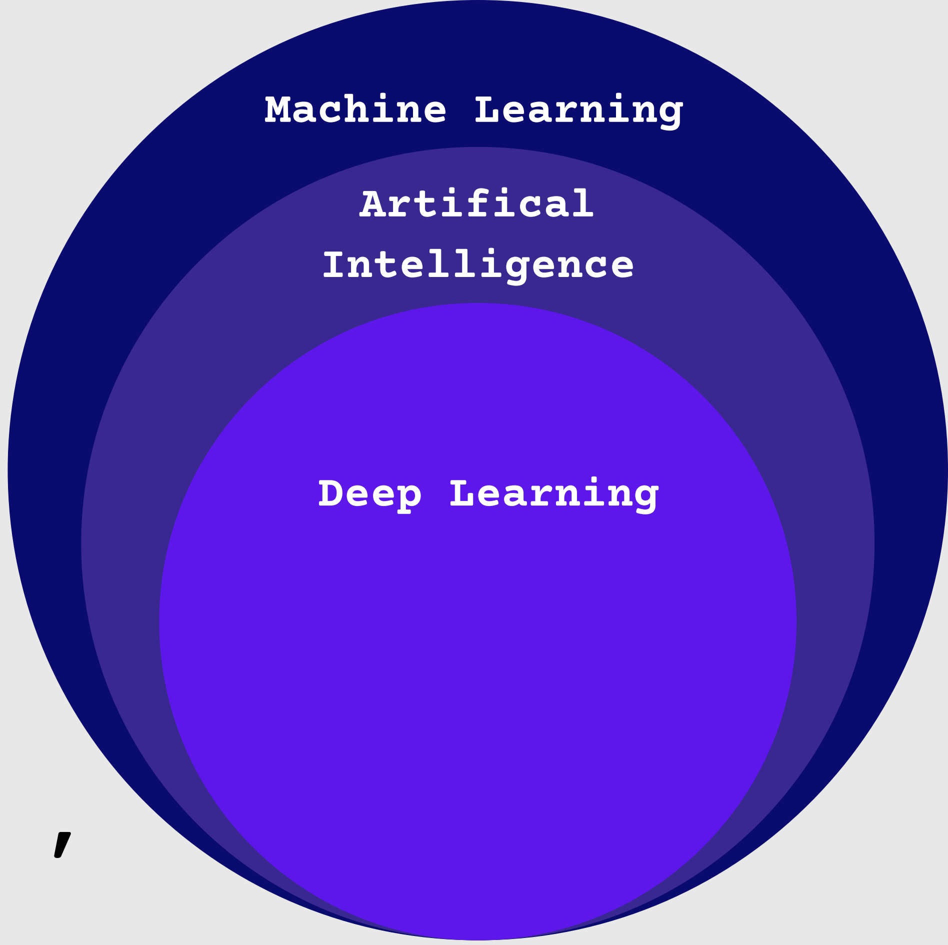
Opening: Erosion followed by dilation(removes small noise).

Closing: Dilation followed by Erosion(closes small holes).

Morphological gradient: Difference between dilation and erosion(detect edges).

What is Deep Learning?

Deep learning is a machine learning technique that uses multi-layered neural networks to automatically learn complex patterns from data.



Neural Networks

Neural networks are inspired by the human brain , consisting of interconnected nodes (neurons) organized in layers. They are widely used for pattern recognition, image classification, and deep learning applications.

Perceptron: The Heart of Neural Networks

A Perceptron is the simplest type of artificial neural network, acting as a binary classifier by processing inputs through weighted sums and an activation function. It forms the foundation of more complex neural networks but is limited to solving only linearly separable problems.

What is an Activation Function?

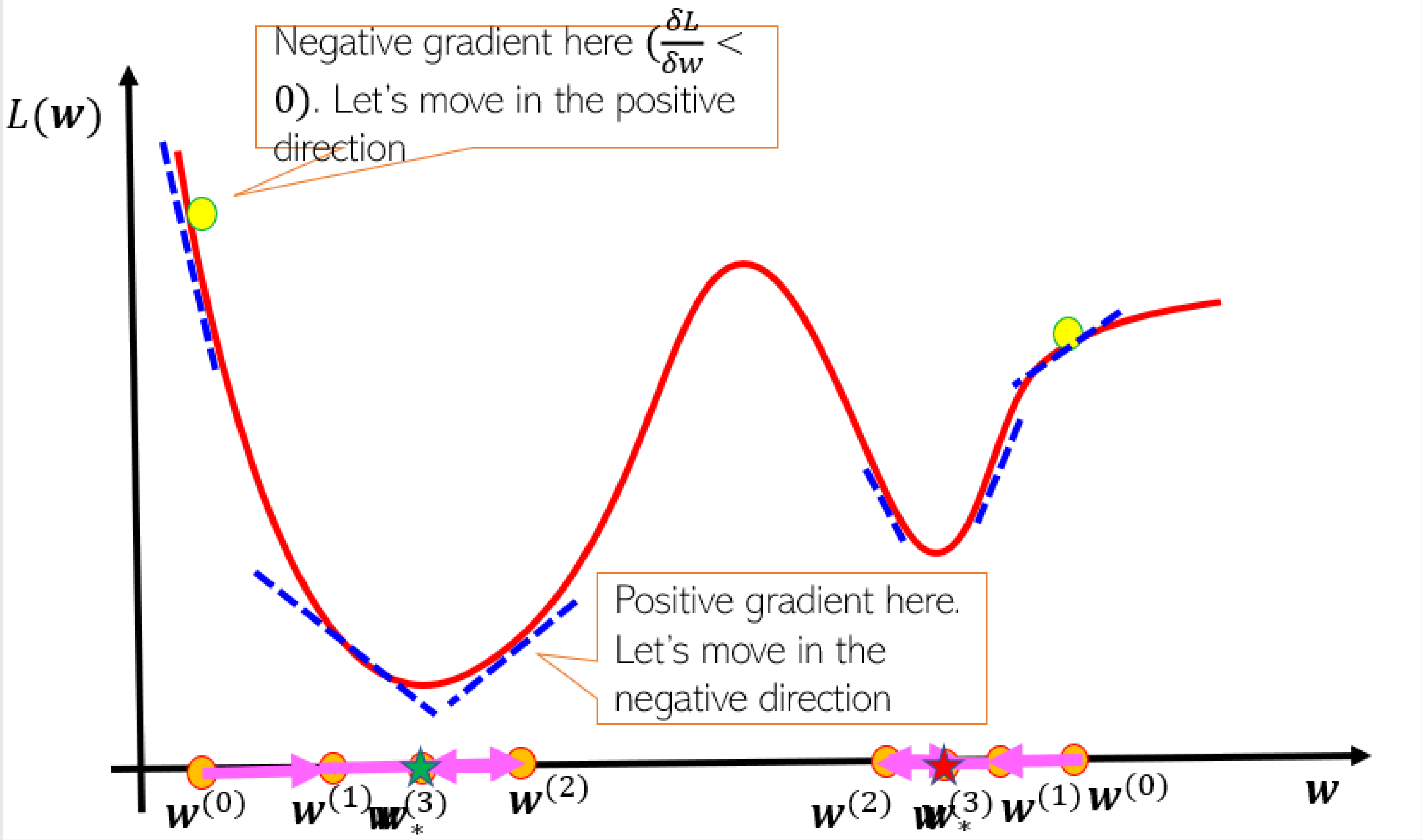
The activation function is a mathematical function that is used within neural networks and decides whether a neuron is activated or not. In simple terms, the activation function determines the strength of the neuron’s response to the weighted input values.

Activation Functions at a Glance

- Sigmoid - Converts input into a value between 0 and 1, often used for probability-based outputs.
- Tanh - Scales input between -1 and 1, making it useful for centered data.
- ReLU - Outputs zero for negative values and the same value for positives, making it efficient for deep networks.
- Leaky ReLU - Similar to ReLU but allows small values for negatives, ensuring continuous learning.
- Softmax - Converts multiple outputs into probabilities that sum to 1, used for multi-class classification.
- Linear - Outputs the input as it is, commonly used in regression problems.

Gradient Descent Algorithm and cost function

The neural network is defined based on inputs and outputs defined based on weighted sums, now cost function is one layer up in complexity. It takes all weights and biases and present you with a single value representing accuracy. The gradient descent takes in those values of weights and biases and tells what causes fastest change in value of cost function. The aim is to minimise the cost function by ending up at a minima.

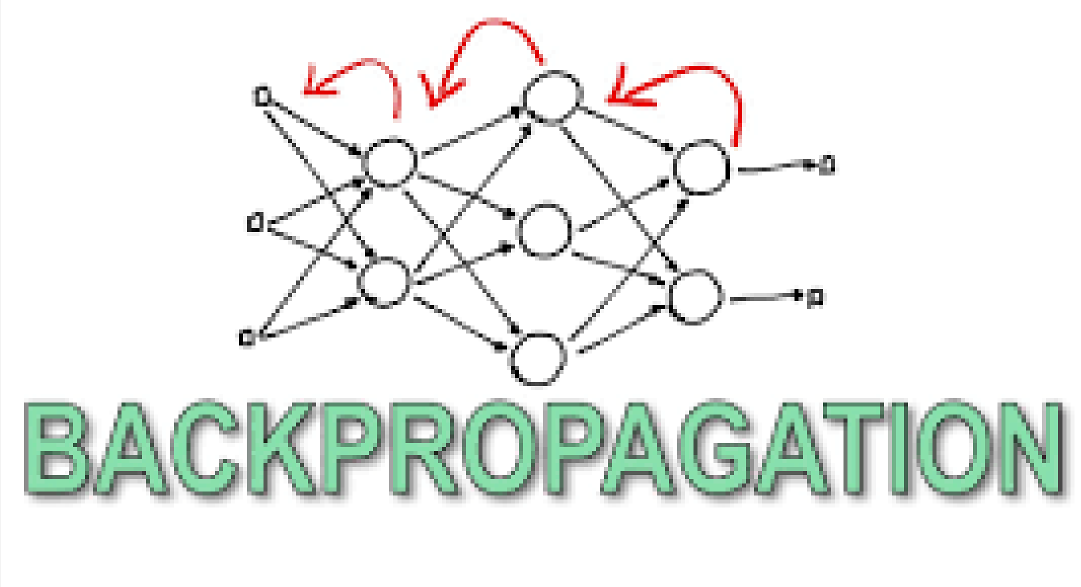


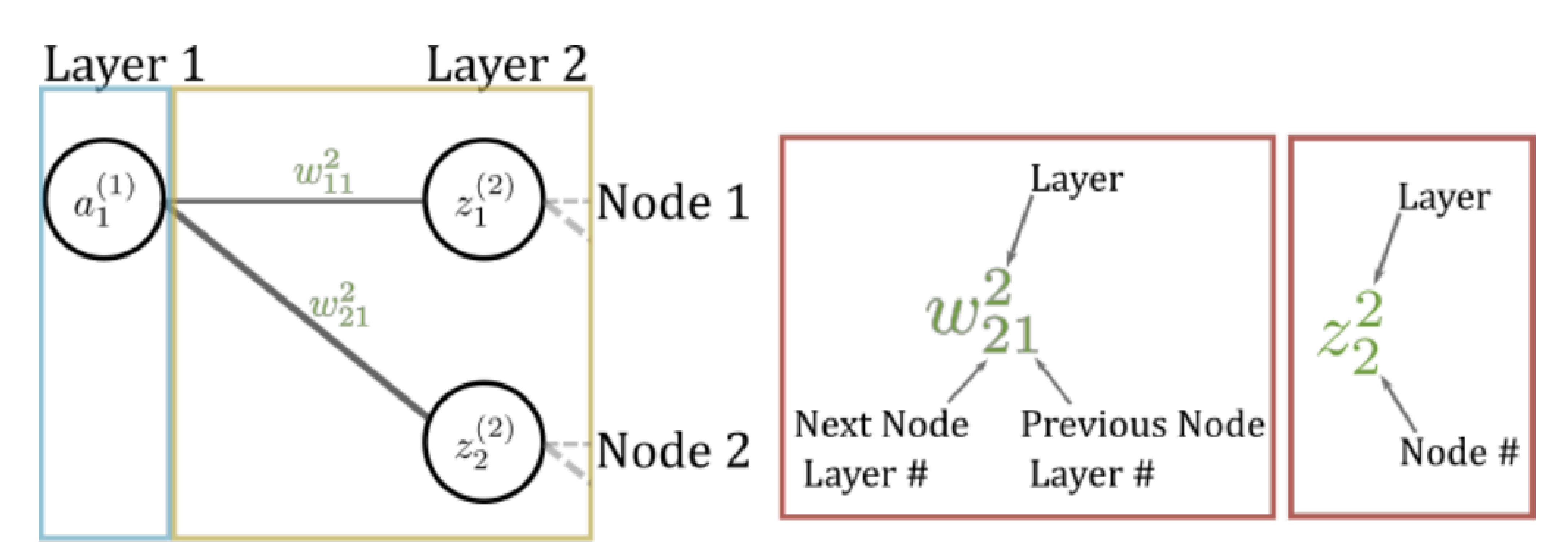
Negative gradient here ($\frac{\delta L}{\delta w} <$ 0). Let's move in the positive direction

Positive gradient here. Let's move in the negative direction

Understanding Backpropagation

Backpropagation in neural networks that adjusts weights to minimize error. It typically involves **maximizing activation** in the forward pass and minimizing error in the backward pass using gradient descent. Like propagating backwards in layers to reduce the cost function.





What are Optimizers in Deep learning?

Optimizers are algorithms that dynamically fine-tune a model's parameters throughout the training process, aiming to minimize a predefined loss function. These specialized algorithms facilitate the learning process of neural networks by iteratively refining the weights and biases based on the feedback received from the data. Well-known optimizers in deep learning encompass Stochastic Gradient Descent (SGD), Adam, and RMSprop.

As a neural network may contain millions of parameters its very crucial to select the right optimizer.

THANK YOU!

Team Entropy