Final Exam (Sample Questions)

1) Otsu's thresholding

Preliminaries:

Your ID sequence will be used to create an image for Problem 1.1 and 1.2. $ID_0 - ID_9$ correspond each of your id as shown below, for example,

6	1	3	1	9	0	6	5	2	1
ID_0	ID_1	ID_2	ID_3	ID_4	ID_5	ID_6	ID_7	ID_8	ID_9

Then, use the quantization table below will be used for quantizing your ID into an intensity value in a range [0, 3]

Original ID (ID_n)	Quantized ID (qID_n)
0-1	0
2-3	1
4-6	2
7-9	3

An 4x10 image created from your quantized ID has 4 different intensity levels in a range of [0,3]

qID_0	qID_1	qID_2	qID_3	qID_4	qID_5	qID_6	qID_7	qID_8	qID_9
qID_0	qID_1	qID_2	qID_3	qID_4	qID_5	qID_6	qID_7	qID_8	qID_9
qID_9	qID_8	qID_7	qID_6	qID_5	qID_4	qID_3	qID_2	qID_1	qID_0
qID_9	qID_8	qID_7	qID_6	qID_5	qID_4	qID_3	qID_2	qID_1	qID_0

Fill in your ID here (Write in the answer sheet provided)

ID_0	ID_1	ID_2	ID_3	ID_4	ID_5	ID_6	ID_7	ID_8	ID_9

Fill in your 4x10 quantized image here

2	0	1	0	3	0	2	2	1	0
2	0	1	0	3	0	2	2	1	0
0	1	2	2	0	3	0	1	0	2
0	1	2	2	0	3	0	1	0	2

Otsu's threshold is a non-parametric and unsupervised method for automated threshold selection. An image has L total levels in a range $[0, \ldots\ L-1]$. To

derive Otsu's threshold, the gray-level histogram is normalized and regarded as a probability distribution:

$$p_i = n_i/N$$

Where n_i is the number of pixels at level i and N is the total number of pixels. A pixel in the image can be classified in two classes, C_0 and C_1 (background and objects) by a threshold level k; C_0 denotes pixels with level $[0, \dots k]$ and C_1 denotes pixels with levels $[k+1, \dots, L-1]$. The probability of class occurrence of C_0 and C_1 denotes $\omega_0\,(k)$ and $\omega_1\,(k)$, respectively, are given by

$$\omega_0\,(k) = \sum_{i=0}^{k} p_i$$

$$\omega_1\,(k) = \sum_{i=k+1}^{L-1} p_i = 1 - \omega_0\,(k)$$

Total mean μ_T is given by

$$\mu_T = \sum_{i=0}^{L-1} ip_i$$

The first-order cumulative moment $\mu(k)$ of the histogram up to level k is givem by

$$\mu(k) = \sum_{i=0}^{k} ip_i$$

The optimal threshold is selected from the value that can maximize the separability of the resultant classes or the between-class variance in the formula below

$$\text{Between-class variance } \sigma_B^2(k) = \frac{(\mu_T\omega_0\,(k) - \mu(k))^2}{\omega_0\,(k)\omega_1\,(k)}$$

Answer the questions 1.1.1-1.1.4 in the answer sheet provided.

1.1.1 Determine total mean (μ_T)

Total mean (μ_T) = 0(0.4) + 1(0.2) + 2(0.3) + 3(0.1) = 1.1

1.1.2 Determine the values in the table

Intensity Level i	**0**	**1**	**2**	**3**
n_i	16	8	12	4
$p_i = n_i/40$	0.4	0.2	0.3	0.1

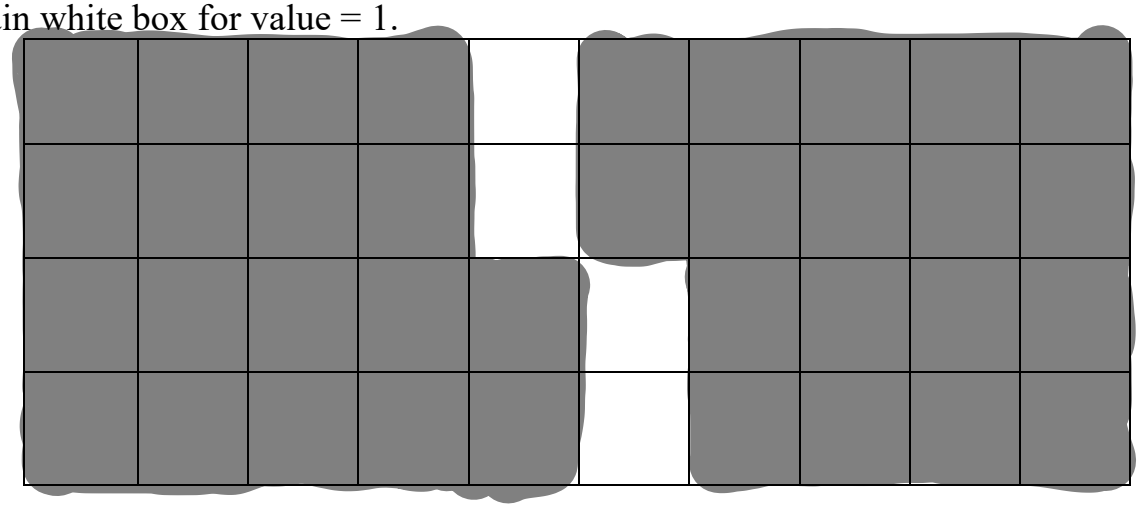
For each selected threshold level k, calculate the between-class variance $\sigma_B^2(k)$ and other parameters below.

Threshold level k	**0**	**1**	**2**	**3**
$\omega_0(k)=\sum_{i=0}^{k} p_i$	0.4	0.4+0.2=0.6	0.4+0.2+0.3 = 0.9	

$\mu(k) = \sum_{i=0}^{k} ip_i$	0	$0(0.4)+1(0.2)$ $= 0.2$	$0.2+2(0.3)$ $= 0.8$	
$\omega_1(k) = 1 - \omega_0(k)$	0.6	0.4	0.1	
Between-Class Variance $\sigma_B^2(k)$	$\frac{(1.1\times0.4-0)^2}{0.6\times0.4}$ $= 0.80667$	$\frac{(1.1\times0.6-0.2)^2}{0.4\times0.6}$ $= 0.88167$	$\frac{(1.1\times0.3-0.8)^2}{(0.1)\times0.9}$ $= 2.4549$	

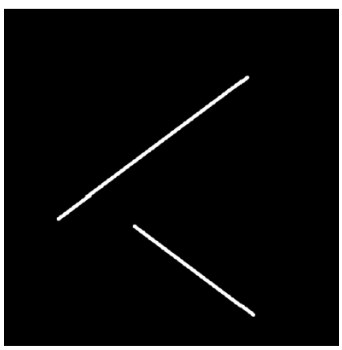
1.1.3 The selected threshold k or Otsu's threshold = 2

1.1.4 Fill in values in the thresholded image after apply inverse thresholding using this Otsu's threshold by **highlighting black color if the value = 0**, and remain white box for value = 1.

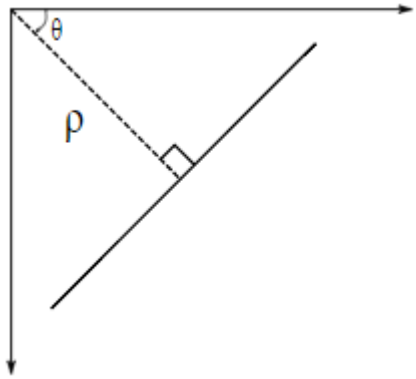


2) Considering the 100 x 100 image (left) with two main lines, draw an output of applying Hough Line transform to the image. The y-axis of the output represents the perpendicular distance (ρ) from the line to the origin (0,0), and the x-axis represents the angle (θ) that this perpendicular makes with the horizontal axis as seen θ-ρ relationship in the right image.

(slide):

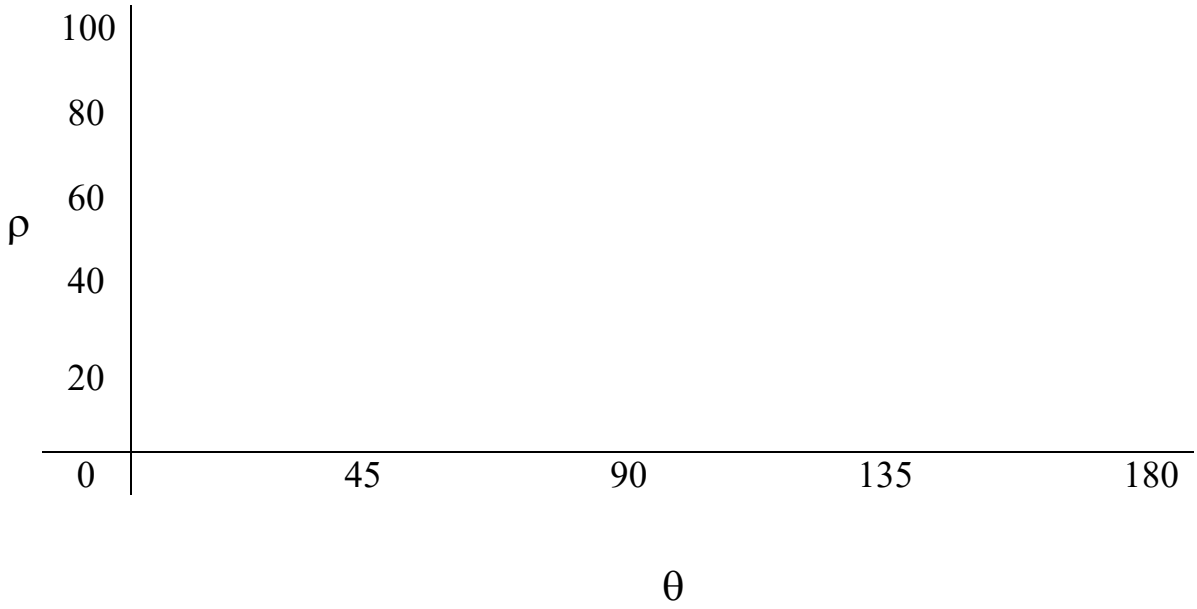


Original Image



θ-ρ relationship

Draw the output below :



Use these options for 3)

A. Erosion
B. Dilation
C. Opening
D. Closing
E. Hit-or-Miss transformation
F. Boundary extraction
G. Region filling
H. Extraction of connected components
I. Convex Hull
J. Thinning
K. Skeletons
L. Pruning

3. Select morphological operations from A - L for the applications in these questions:

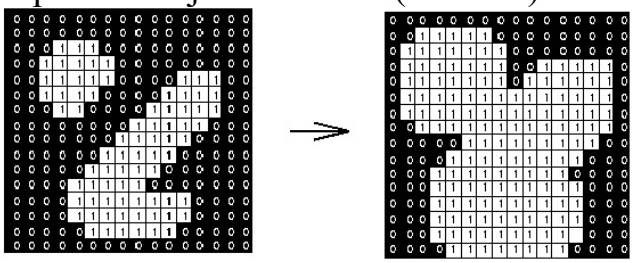
3.1 Find the locations (those pixels) whose neighbourhood matches the shape of a targeted structuring element (2 points)

Ans: E Hit-or-Miss Transformation

3.2 To clean up some noise in the image below, but maintaining the same size of the image. (2 Points)

Ans: C Opening (Erode -> dilate)

3.3 Which morphological technique should be used to merge between two separated objects below? (2 Points)

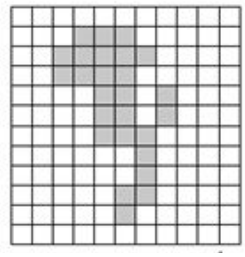


Ans: B. Dilation

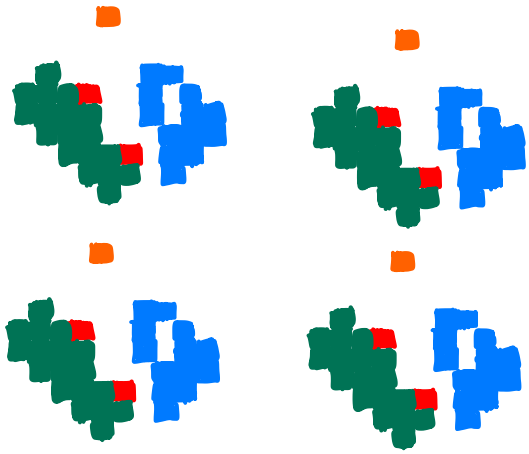
3.4) A method to find object structure in only single-pixel wide lines. (2 points)

Ans: J. or K. Thinning or Skeleton

4. Find the convex hull of a binary shape below



Ans (slide):



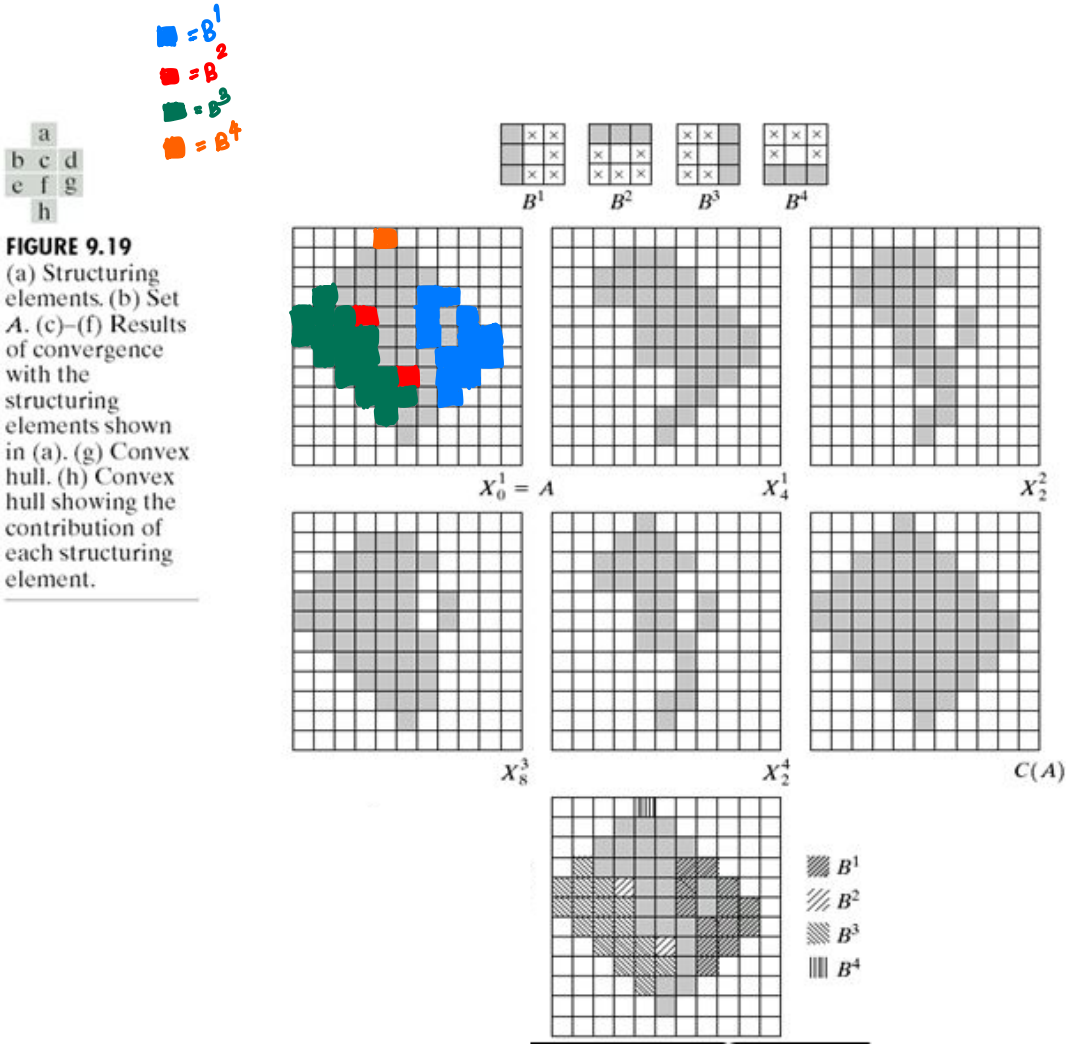


FIGURE 9.19
(a) Structuring elements. (b) Set A. (c)–(f) Results of convergence with the structuring elements shown in (a). (g) Convex hull. (h) Convex hull showing the contribution of each structuring element.

5. Convolution Neural Networks

5.1) What is the purpose of Convolution? (2 points)

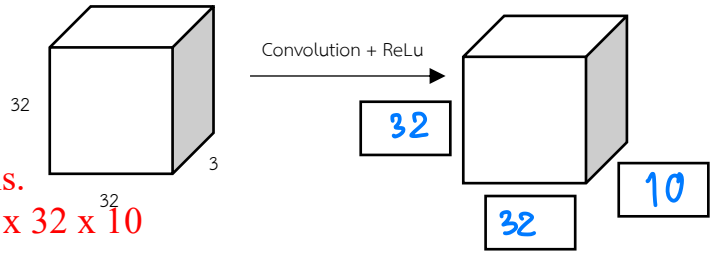
Ans. Extract features from the image using learnable filters.

5.2) What is max-pooling? (2 points)

Ans. The process to reduce dimension of the feature size by selecting the max value from each kernel.

5.3) Fill the size of the output in the blank box below when convolve the image (32 x 32 x 3) by 10 filters (3 x 3 x 3) with stride = 1, with 1 pixel padding (3 points)

$$\frac{\text{width} + 2\text{padding} - \text{filter\_size}}{\text{stride}} + 1$$



$$\frac{32 + 2(1) - 3}{1} + 1$$

$$= 31 + 1 = 32$$

Ans.
32 x 32 x 10

5.4) From 5.3) How many parameters (weights and bias) for convolutional layer below? (3 points)
Ans.
(32 x 32 x 3 +1) * 10 = 30730 parameters

6. Describe the diffusion process

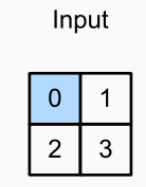
7. Variants of modern models
7.1) What is a key advantage of Vision Transformers (ViT) compared to traditional convolutional neural networks?

7.2) What is the main innovation introduced by ResNet (Residual Networks)?

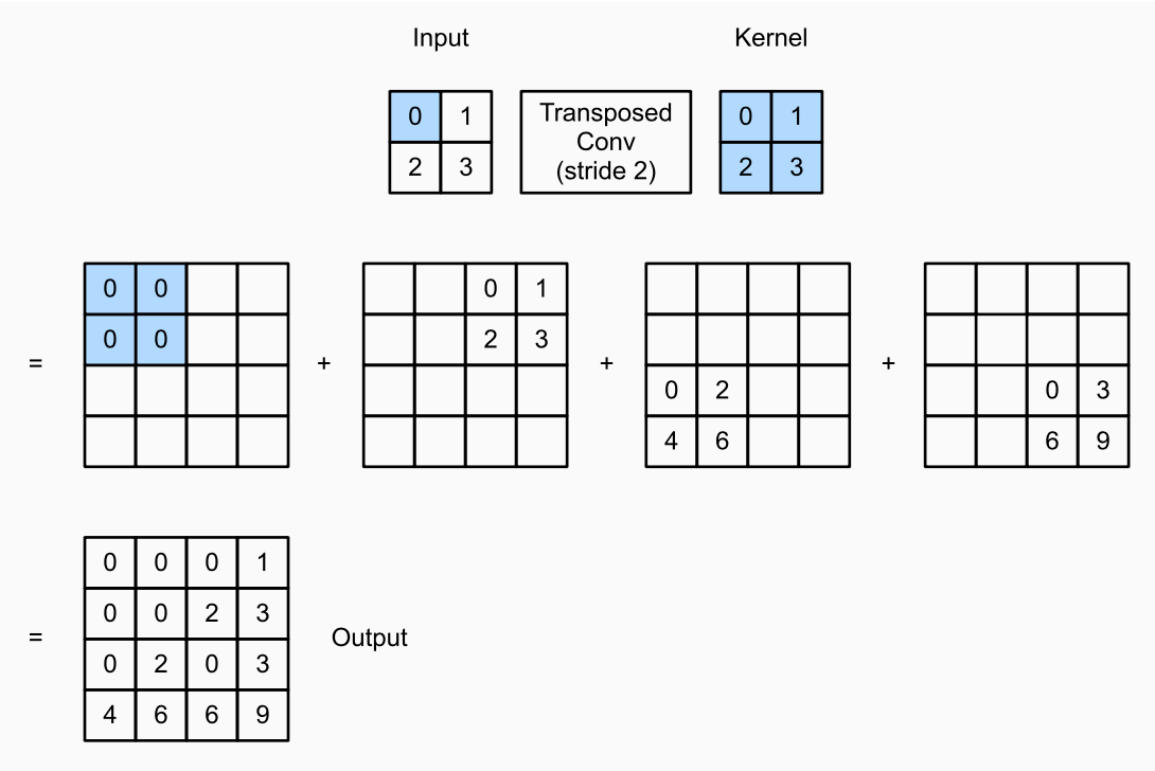
7.3) What is the key characteristic of EfficientNet's scaling strategy?

8) Classification application
Design a classification model for recognizing 20 kinds of Thai fruits, for example, banana, grape, watermelon, coconut, etc. Describe the process and write a flow diagram from collecting datasets, pre-processing, modelling, parameters, and evaluation method used in the application (10 points).

9. Deconvolution / Transposed Convolution
Determine the output from Transposed convolution using a 2 x 2 kernel, stride of 2, no padding of the image below:



Ans (slide):



10. U-Net

10.1 Design 4 convolutional-layer encoder-decoder of U-Net using 3x3 convolution without padding, 2x2 max-pooling. Each layer has two-conv layer. (The actual exam will identify components required)

Ans (slide):

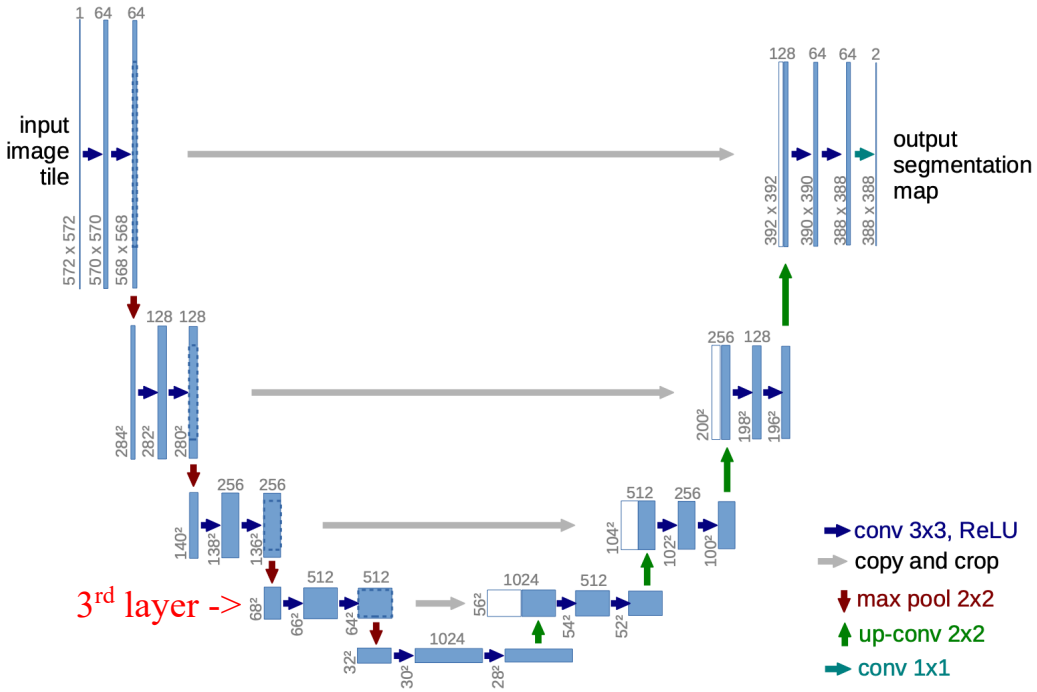


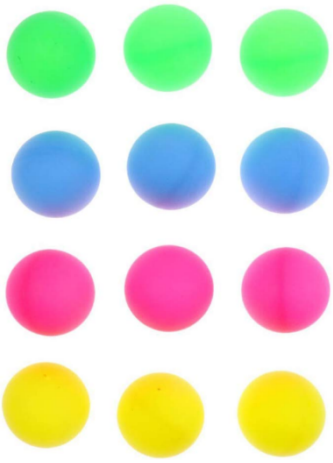
Fig. 1. U-net architecture (example for 32x32 pixels in the lowest resolution). Each blue box corresponds to a multi-channel feature map. The number of channels is denoted on top of the box. The x-y-size is provided at the lower left edge of the box. White boxes represent copied feature maps. The arrows denote the different operations.

10.2 Calculate the number of learnable parameters in the 3rd layer of the encoder (with bias).

11. Design a Generative Adversarial Network (GAN) for an image-to-image translation task, such as translating grayscale images to colored images or transforming photos into artistic styles. Your task is to:

- Draw a schematic diagram of the GAN architecture, clearly showing the generator and discriminator.
- Label the key components in the generator and describe their role.
- Label the key components in the discriminator and describe their role.
- Briefly explain the loss functions.

13. Design a process using image processing to count the number of ping pong balls of different colors as shown in the image below. How are you going to evaluate your method. Also, describe the constraints of your selected approach.

Image	Output
	number of the yellow pingpong balls = 3 number of the pink pingpong balls = 3 number of the blue pingpong balls = 3 number of the green pingpong balls = 3
	number of the yellow pingpong balls = 3 number of the pink pingpong balls = 3 number of the blue pingpong balls = 3 number of the green pingpong balls = 3
	number of the yellow pingpong balls = 3 number of the pink pingpong balls = 3 number of the blue pingpong balls = 3 number of the green pingpong balls = 3