
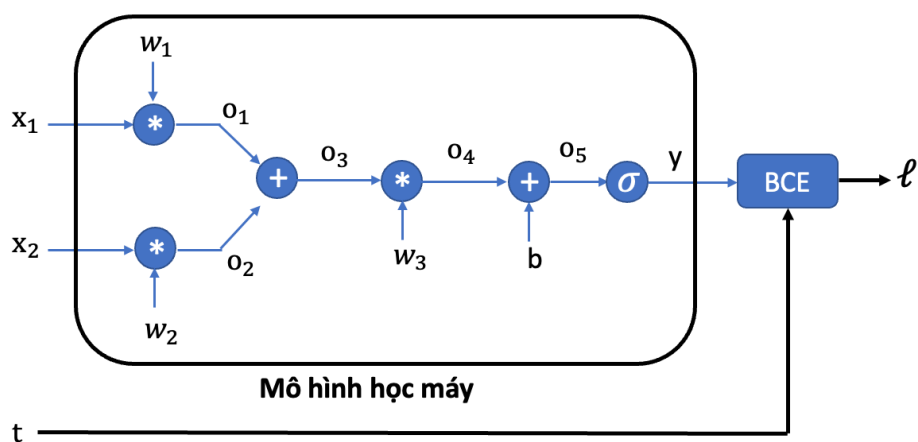


Giảng viên ra đề: (Chữ ký và Họ tên) Lê Thành Sách	(Ngày ra đề) 29/11/2022	Người phê duyệt: (Chữ ký, Chức vụ và Họ tên)	(Ngày duyệt đề)
(phần phía trên cần che đi khi in sao đề thi)			

 TRƯỜNG ĐẠI HỌC BÁCH KHOA – ĐHQG-HCM KHOA KH & KT MÁY TÍNH	THI CUỐI KỲ		Học kỳ/năm học		1	2022-2023
			Ngày thi		29/11/2022	
	Môn học	Học sâu và ứng dụng trong thị giác máy tính				
	Mã môn học	055255				
	Thời lượng	110 phút	Mã đề	0001		
Ghi chú: <div>- Được sử dụng tài liệu giấy - Không được trao đổi thông tin với người khác bằng bất cứ hình thức nào - Nộp lại đề cùng với giấy bài làm</div>						

Câu 1. (L.O.1,2) (2 điểm): Cho mô hình học máy, có hai đầu vào và một đầu ra, như trong Hình 1.



- σ : Hàm sigmoid, được định nghĩa và có đạo hàm như sau:

$$y = \sigma(z) = \frac{1}{1 + e^{-z}} \quad y' = y(1 - y)$$

- BCE: Binary CrossEntropy, được định nghĩa và có đạo hàm như sau:

$$\ell = \text{BCE}(y, t) = -[t \log(y) + (1 - t) \log(1 - y)] \quad \frac{d\ell}{dy} = \frac{y - t}{y(1 - y)}$$

Hình 1. Mô hình học máy và sơ đồ huấn luyện

Bảng 1. Tập huấn luyện

STT	x_1	x_2	t
1	-2	0	1
2	2	1	0

- a) **[1,0đ]** Hãy cập nhật các tham số của mô hình, dùng giải thuật gradient descent. Cho biết: (i) Giá trị hiện tại của các tham số là: $w_1 = +1$; $w_2 = -1$; $w_3 = +1$; $b = 1$; (ii) hệ số học: $\alpha = 0.01$; batch-size=1; dùng 01 epoch.

• Batch-1: $\{< (-2, 0), 1 >\}$

○ Forward:

- $o_1 = x_1 * w_1 = -2 * 1 = -2; \frac{\partial o_1}{\partial w_1} = x_1 = -2$
- $o_2 = x_2 * w_2 = 0 * -1 = 0; \frac{\partial o_2}{\partial w_2} = x_2 = 0$
- $o_3 = o_1 + o_2 = -2 + 0 = -2; \frac{\partial o_3}{\partial o_1} = 1; \frac{\partial o_3}{\partial o_2} = +1$
- $o_4 = o_3 * w_3 = -2 * 1 = -2; \frac{\partial o_4}{\partial w_3} = o_3 = -2; \frac{\partial o_4}{\partial o_3} = w_3 = +1$
- $o_5 = o_4 + b = -2 + 1 = -1; \frac{\partial o_5}{\partial b} = 1; \frac{\partial o_5}{\partial o_4} = +1$
- $y = \sigma(o_5) = \frac{1}{1+e^1} = 0.269; \frac{\partial y}{\partial o_5} = y(1-y) = 0.269 * (1 - 0.269) = 0.197$
- $\ell = BCE(y, t) = -\log(y) = -\log(0.269) = 1.313 \text{ (base } e\text{)};$
- $\frac{\partial \ell}{\partial o_5} = \frac{\partial \ell}{\partial y} \times \frac{\partial y}{\partial o_5} = y - t = 0.269 - 1 = -0.731$

○ Backward:

- $\Delta w_1 = \frac{\partial o_1}{\partial w_1} \times \frac{\partial o_3}{\partial o_1} \times \frac{\partial o_4}{\partial o_3} \times \frac{\partial o_5}{\partial o_4} \times \frac{\partial \ell}{\partial o_5}$
 $= (-2) \times 1 \times 1 \times 1 \times (-0.731) = 1.462$
- $\Delta w_2 = \frac{\partial o_2}{\partial w_2} \times \frac{\partial o_3}{\partial o_2} \times \frac{\partial o_4}{\partial o_3} \times \frac{\partial o_5}{\partial o_4} \times \frac{\partial \ell}{\partial o_5}$
 $= (0) \times 1 \times 1 \times 1 \times (-0.731) = 0$
- $\Delta w_3 = \frac{\partial o_4}{\partial w_3} \times \frac{\partial o_5}{\partial o_4} \times \frac{\partial \ell}{\partial o_5} = (-2) \times 1 \times (-0.731) = 1.462$
- $\Delta b = \frac{\partial o_5}{\partial b} \times \frac{\partial \ell}{\partial o_5} = 1 \times (-0.731) = -0.731$

○ Update:

- $w_1 = w_1 - \alpha \times \Delta w_1 = 1 - 0.01 \times 1.462 = 0.985$
- $w_2 = w_2 - \alpha \times \Delta w_2 = -1 - 0.01 \times 0 = -1$
- $w_3 = w_3 - \alpha \times \Delta w_3 = 1 - 0.01 \times 1.462 = 0.985$
- $b = b - \alpha \times \Delta b = 1 - 0.01 \times (-0.731) = 1.007$

• Batch-2: $\{< (2, 1), 0 >\}$

○ Forward:

- $o_1 = x_1 * w_1 = 2 * 0.985 = 1.97; \frac{\partial o_1}{\partial w_1} = x_1 = +2$
- $o_2 = x_2 * w_2 = 1 * -1 = -1; \frac{\partial o_2}{\partial w_2} = x_2 = 1$
- $o_3 = o_1 + o_2 = 1.97 - 1 = 0.97; \frac{\partial o_3}{\partial o_1} = 1; \frac{\partial o_3}{\partial o_2} = +1$
- $o_4 = o_3 * w_3 = 0.97 * 0.985 = 0.955; \frac{\partial o_4}{\partial w_3} = o_3 = 0.97; \frac{\partial o_4}{\partial o_3} = w_3 = 0.985$
- $o_5 = o_4 + b = 0.955 + 1.007 = 1.962; \frac{\partial o_5}{\partial b} = 1; \frac{\partial o_5}{\partial o_4} = +1$
- $y = \sigma(o_5) = \frac{1}{1+e^{-1.962}} = 0.877; \frac{\partial y}{\partial o_5} = y(1-y) = 0.877 * (1 - 0.877) = 0.108$
- $\ell = BCE(y, t) = -\log(y) = -\log(0.877) = 0.131 \text{ (base } e\text{)};$
- $\frac{\partial \ell}{\partial o_5} = \frac{\partial \ell}{\partial y} \times \frac{\partial y}{\partial o_5} = y - t = 0.877 - 0 = 0.877$

○ Backward:

- $\Delta w_1 = \frac{\partial o_1}{\partial w_1} \times \frac{\partial o_3}{\partial o_1} \times \frac{\partial o_4}{\partial o_3} \times \frac{\partial o_5}{\partial o_4} \times \frac{\partial \ell}{\partial o_5}$
 $= 2 \times 1 \times 0.985 \times 1 \times 0.877 = 1.728$
- $\Delta w_2 = \frac{\partial o_2}{\partial w_2} \times \frac{\partial o_3}{\partial o_2} \times \frac{\partial o_4}{\partial o_3} \times \frac{\partial o_5}{\partial o_4} \times \frac{\partial \ell}{\partial o_5}$

$$= 1 \times 1 \times 0.985 \times 1 \times 0.877 = 0.864$$

$$\begin{aligned} \Delta w_3 &= \frac{\partial o_4}{\partial w_3} \times \frac{\partial o_5}{\partial o_4} \times \frac{\partial \ell}{\partial o_5} = 0.97 \times 1 \times 0.877 = 0.851 \\ \Delta b &= \frac{\partial o_5}{\partial b} \times \frac{\partial \ell}{\partial o_5} = 1 \times 0.877 = 0.877 \end{aligned}$$

○ Update:

$$\begin{aligned} w_1 &= w_1 - \alpha \times \Delta w_1 = 0.985 - 0.01 \times 1.728 = 0.968 \\ w_2 &= w_2 - \alpha \times \Delta w_2 = -1 - 0.01 \times 0.864 = -1.009 \\ w_3 &= w_3 - \alpha \times \Delta w_3 = 0.985 - 0.01 \times 0.851 = 0.976 \\ b &= b - \alpha \times \Delta b = 1.007 - 0.01 \times 0.877 = 0.998 \end{aligned}$$

b) **[0,5đ]** Các nhãn của bài toán là: $\{1 \rightarrow \text{Bệnh}; 0 \rightarrow \text{Khỏe}\}$; Xác định nhãn cho các điểm dữ liệu trong tập huấn luyện trước và sau khi cập nhật ở Câu (a). Cho biết giá trị ngưỡng để xác định nhãn là 0.5

▪ Trước khi huấn luyện: $w_1 = +1$; $w_2 = -1$; $w_3 = +1$; $b = 1$

▪ Mẫu: $(-2, 0)$

$$\begin{aligned} o_1 &= x_1 * w_1 = -2 * 1 = -2; \\ o_2 &= x_2 * w_2 = 0 * -1 = 0; \\ o_3 &= o_1 + o_2 = -2 + 0 = -2; \\ o_4 &= o_3 * w_3 = -2 * 1 = -2; \\ o_5 &= o_4 + b = -2 + 1 = -1; \\ y &= \sigma(o_5) = \frac{1}{1+e^1} = 0.269 \end{aligned}$$

○ Nhãn: $y < 0.5 \Rightarrow 0$ (khỏe)

▪ Mẫu: $(2, 1)$

$$\begin{aligned} o_1 &= x_1 * w_1 = 2 * 1 = 2; \\ o_2 &= x_2 * w_2 = 1 * -1 = -1; \\ o_3 &= o_1 + o_2 = 2 - 1 = 1; \\ o_4 &= o_3 * w_3 = 1 * 1 = 1; \\ o_5 &= o_4 + b = 1 + 1 = +2; \\ y &= \sigma(o_5) = \frac{1}{1+e^{-2}} = 0.881 \end{aligned}$$

○ Nhãn: $y > 0.5 \Rightarrow 1$ (bệnh)

▪ Sau khi huấn luyện: $w_1 = 0.968$; $w_2 = -1.009$; $w_3 = 0.976$; $b = 0.998$

▪ Mẫu: $(-2, 0)$

$$\begin{aligned} o_1 &= x_1 * w_1 = -2 * 0.968 = -1.936; \\ o_2 &= x_2 * w_2 = 0 * -1.009 = 0; \\ o_3 &= o_1 + o_2 = -1.936 + 0 = -1.936; \\ o_4 &= o_3 * w_3 = -1.936 * 0.976 = -1.890; \\ o_5 &= o_4 + b = -1.890 + 0.998 = -0.892; \\ y &= \sigma(o_5) = \frac{1}{1+e^{0.892}} = 0.291 \end{aligned}$$

○ Nhãn: $y < 0.5 \Rightarrow 0$ (khỏe)

▪ Mẫu: $(2, 1)$

$$\begin{aligned} o_1 &= x_1 * w_1 = 2 * 0.968 = 1.936; \\ o_2 &= x_2 * w_2 = 1 * -1.009 = -1.009; \\ o_3 &= o_1 + o_2 = 1.936 - 1.009 = 0.927; \\ o_4 &= o_3 * w_3 = 0.927 * 0.976 = 0.905; \\ o_5 &= o_4 + b = 0.905 + 0.998 = 1.903; \\ y &= \sigma(o_5) = \frac{1}{1+e^{-1.903}} = 1.0 \end{aligned}$$

○ Nhãn: $y > 0.5 \Rightarrow 1$ (bệnh)

c) [0,5đ] Điều gì xảy ra trong lúc huấn luyện, nếu ràng buộc giá trị tham số w_3 trong miền có giá trị: (i) rất nhỏ gần 0; và (ii) trong miền có giá trị rất lớn?

• Ta có: $\frac{\partial o_4}{\partial o_3} = w_3$

○ Nếu w_3 nhỏ (gần 0) thì

▪ $\Delta w_1 = \frac{\partial o_1}{\partial w_1} \times \frac{\partial o_3}{\partial o_1} \times \frac{\partial o_4}{\partial o_3} \times \frac{\partial o_5}{\partial o_4} \times \frac{\partial \ell}{\partial o_5}$, và

▪ $\Delta w_2 = \frac{\partial o_2}{\partial w_2} \times \frac{\partial o_3}{\partial o_2} \times \frac{\partial o_4}{\partial o_3} \times \frac{\partial o_5}{\partial o_4} \times \frac{\partial \ell}{\partial o_5}$ rất nhỏ

▪ Do đó, các tham số w_1 và w_2 hầu như không thay đổi trong quá trình học

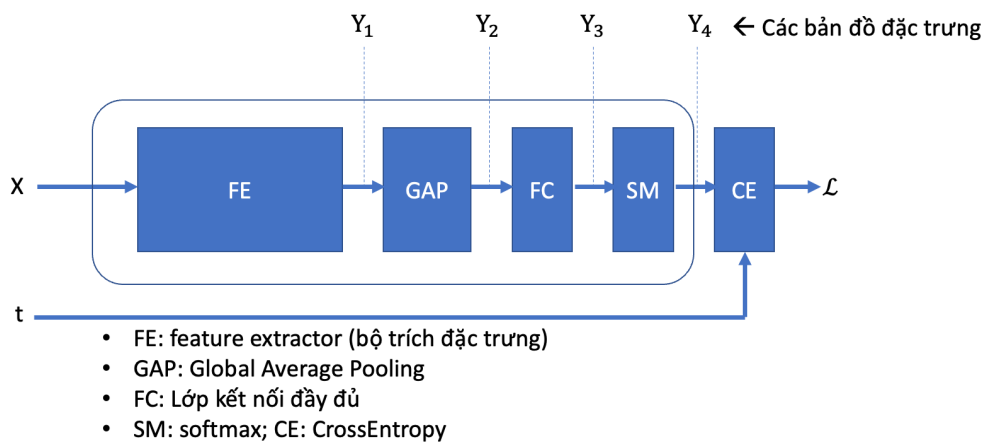
○ Nếu w_3 rất lớn thì

▪ $\Delta w_1 = \frac{\partial o_1}{\partial w_1} \times \frac{\partial o_3}{\partial o_1} \times \frac{\partial o_4}{\partial o_3} \times \frac{\partial o_5}{\partial o_4} \times \frac{\partial \ell}{\partial o_5}$, và

▪ $\Delta w_2 = \frac{\partial o_2}{\partial w_2} \times \frac{\partial o_3}{\partial o_2} \times \frac{\partial o_4}{\partial o_3} \times \frac{\partial o_5}{\partial o_4} \times \frac{\partial \ell}{\partial o_5}$ có thể giá trị rất lớn (khi các đạo hàm khác không quá nhỏ)

▪ Do đó, các tham số w_1 và w_2 có sự thay đổi với biên độ lớn và mô hình không ổn định.

Câu 2. (L.O.1,2,3) (2 điểm): Cho mô hình học máy như trong Hình 2, dùng để phân loại với 3 nhãn: {0→Chó; 1→Mèo; 2→Gà}



Hình 2. Sơ đồ huấn luyện

a) [1,0đ] Xác định số tham số cần phải học của mô hình và vùng ảnh hưởng cho mỗi đặc trưng của Y_1 trong ảnh đầu vào, với thông tin cho như sau:

- FE là chuỗi của các phép tính toán như sau:
 - 128 bộ tích chập (conv) → 64 bộ tích chập (conv) → Maxpool → 4 bộ tích chập (conv)
 - Tất cả các bộ tích chập: kích thước 3x3; no padding; stride=1; bias=True; activation=relu; Maxpool: 2x2, stride=2
- FC: Kết nối đầy đủ, Bias = True.

Số lượng tham số của mô hình: giả sử ảnh đầu vào 3 kênh màu:

- Khối FE:
 - 128 bộ tích chập (conv): $128 \times 3 \times 3 \times 3 + 128 = 3.584$
 - 64 bộ tích chập (conv) : $64 \times 3 \times 3 \times 128 + 64 = 7.3792$
 - 4 bộ tích chập (conv) : $4 \times 3 \times 3 \times 64 + 4 = 2.308$
- Khối FC: $4 \times 3 + 3 = 15$
- Tổng cộng: $3.584 + 73.792 + 2.308 + 15 = 76.699$

Vùng ảnh hưởng của mỗi đặc trưng của Y_1 : 10×10 , xem giải thích sau

Input \rightarrow 128 conv \rightarrow 64 conv \rightarrow Maxpool \rightarrow 4 conv \rightarrow (FE-out)
 10×10 8×8 6×6 3×3 1

- b) **[0,5đ]** Tính toán đáp ứng của mô hình và xác định nhãn cho các điểm dữ liệu trong Bảng 2, chỉ dùng cột Y_3 .

Bảng 2. Bảng dữ liệu

STT	x	Nhãn	Y_3 (3 hàng tương ứng cho: x_1, x_2, x_3)
1	x_1	“Chó”	[[1000, 1001, 1009], [20, 15, 10], [45, 46, 44]]
2	x_2	“Mèo”	
3	x_3	“Gà”	

- $\text{softmax}(Y_3)$:
 $[[1.23353201\text{e-}04, 3.35308764\text{e-}04, \mathbf{9.99541338\text{e-}01}],$
 $[\mathbf{9.93262357\text{e-}01}, 6.69254912\text{e-}03, 4.50940412\text{e-}05],$
 $[2.44728471\text{e-}01, \mathbf{6.65240956\text{e-}01}, 9.00305732\text{e-}02]]$
- $\text{argmax}(Y_4, \text{axis} = -1) \Rightarrow$ nhãn: [2, 0, 1] hay [Gà, Chó, Mèo]

- c) **[0,5đ]** Xác định hình dạng của các bản đồ: Y_1, Y_2, Y_3, Y_4 nếu batch dữ liệu đầu vào là: (10, 224, 224, 3); lưu ý: sử dụng quy ước channel-last.

Hình dạng dữ liệu:

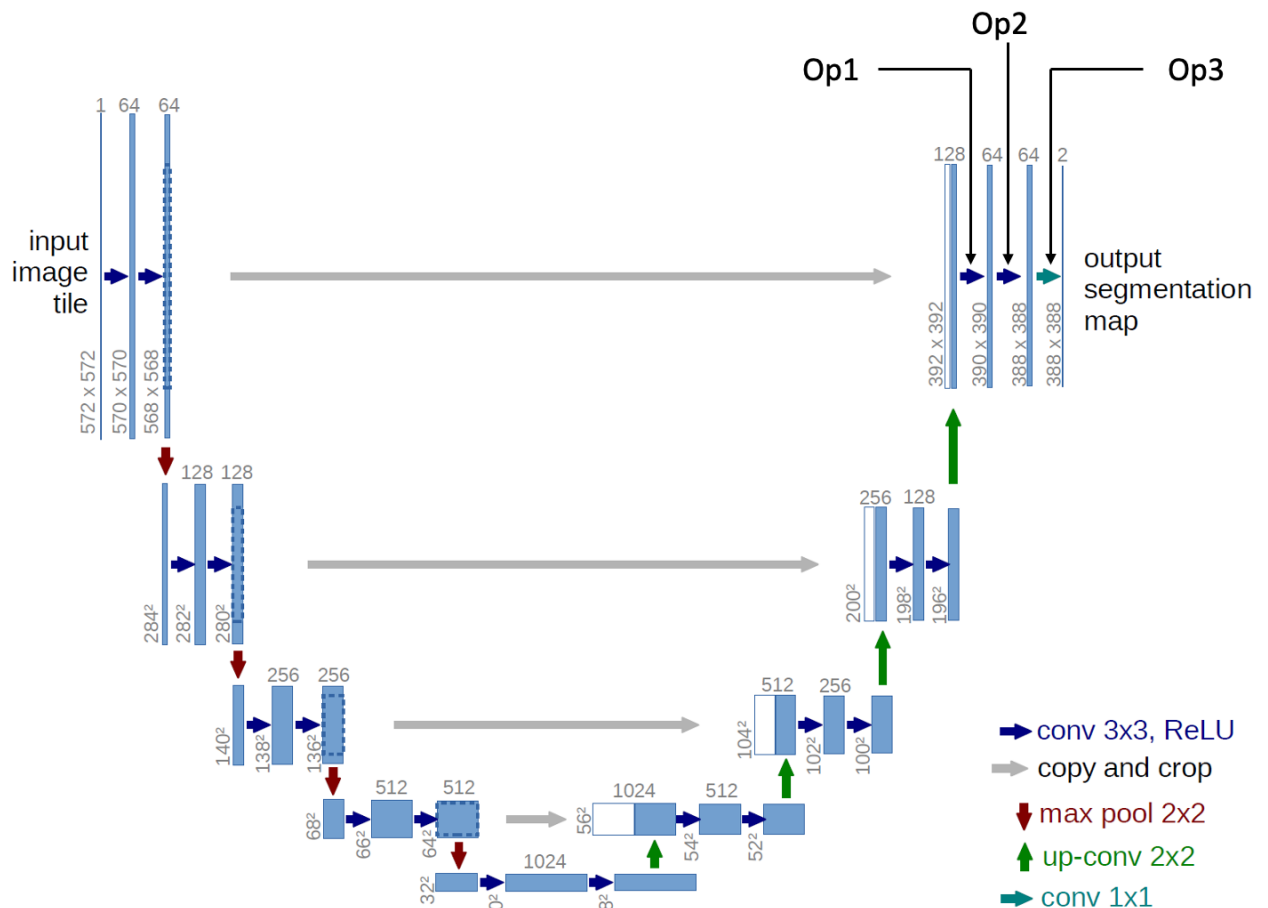
- Khối FE:

Phép toán	Hình dạng kết quả	Tên gọi
Input	(10, 224, 224, 3)	Input
128 conv	(10, 222, 222, 128)	
64 conv	(10, 220, 220, 64)	
Maxpool	(10, 110, 110, 64)	
4 conv	(10, 108, 108, 4)	Y1
GAP	(10, 4)	Y2
FC	(10,3)	Y3

SM	(10,3)	Y4
----	--------	----

Câu 3. (L.O. 1,2,3) (2 điểm): Với mô hình Unet như trong Hình 3.

- a) **[0,5đ]** Xác định số tham số cần học của nhánh Encoder (gồm các phép tính toán phía trước phép **up-conv** đầu tiên trong đồ thị tính toán. Cho biết: tất cả các conv là có dùng bias; kích thước =3x3, stride=1.
- b) **[0,5đ]** Gọi Op1, Op2, và Op3 là 3 phép toán sau cùng trên đồ thị; hãy xác định số lượng phép tích chập (số bộ lọc) trong từng phép toán đó và hàm activation của phép Op3?



Hình 3. Mô hình Unet

STT	Phép toán	Hình dạng filters	Hình dạng kết quả	# Tham số
1	Input		(N, 572, 572, 1)	
2	Conv1-1	(64, 3, 3, 1)	(N, 570, 570, 64)	$64 \times 3 \times 3 \times 1 + 64 = 640$
3	Conv1-2	(64, 3, 3, 64)	(N, 568, 568, 64)	$64 \times 3 \times 3 \times 64 + 64 = 36.928$
4	Maxpool		(N, 284, 284, 64)	
5	Conv2-1	(128, 3, 3, 64)	(N, 282, 282, 128)	$128 \times 3 \times 3 \times 64 + 128 = 73.856$

6	Conv2-2	(128,3,3,128)	(N, 280, 280, 128)	$128 \times 3 \times 3 \times 128 + 128 = 147.584$
7	Maxpool		(N, 140, 140, 128)	
8	Conv3-1	(256, 3, 3, 128)	(N, 138, 138, 256)	$256 \times 3 \times 3 \times 128 + 256 = 295.168$
9	Conv3-2	(256, 3, 3, 256)	(N, 136, 136, 256)	$256 \times 3 \times 3 \times 256 + 256 = 590.080$
10	Maxpool		(N, 68, 68, 256)	
11	Conv4-1	(512,3,3,256)	(N, 66, 66, 512)	$512 \times 3 \times 3 \times 256 + 512 = 1.180.160$
12	Conv4-2	(512,3,3,512)	(N, 64, 64, 512)	$512 \times 3 \times 3 \times 512 + 512 = 2.359.808$
13	Maxpool		(N, 32, 32, 512)	
14	Conv5-1	(1024,3,3,512)	(N, 30, 30, 1024)	$1024 \times 3 \times 3 \times 512 + 1024 = 4.719.616$
15	Conv5-2	(512,3,3,1024)	(N, 28, 28, 512)	$512 \times 3 \times 3 \times 1024 + 512 = 4.719.104$
			Tổng cộng	

c) **[0,5đ]** Bài toán phân đoạn của mô hình có mấy nhãn? Vì sao?

- Số nhãn của mô hình: 2
- Lý do: đầu ra của Op3 là (N, 388,388,2); nghĩa là mỗi pixel có phân phối xác suất cho hai nhãn (tổng bằng 1); và để xác định xác định nhãn có thể dùng $\text{argmax}(y, \text{axis}=-1)$; với y là đầu ra của Op3

d) **[0,5đ]** Gọi K là số nhãn của bài toán phân đoạn; viết code (Keras hay Pytorch) để mô tả *phép toán Op3*. Lưu ý, Op3 đã có hàm activation bên trong, do đó chú ý viết code cho đúng với thư viện Pytorch

- **Keras:**

```
import keras.layers as layers
Op3 = layers.Conv2D(K, kernel-size=1, stride=(1,1), padding="same", activation="softmax")
```

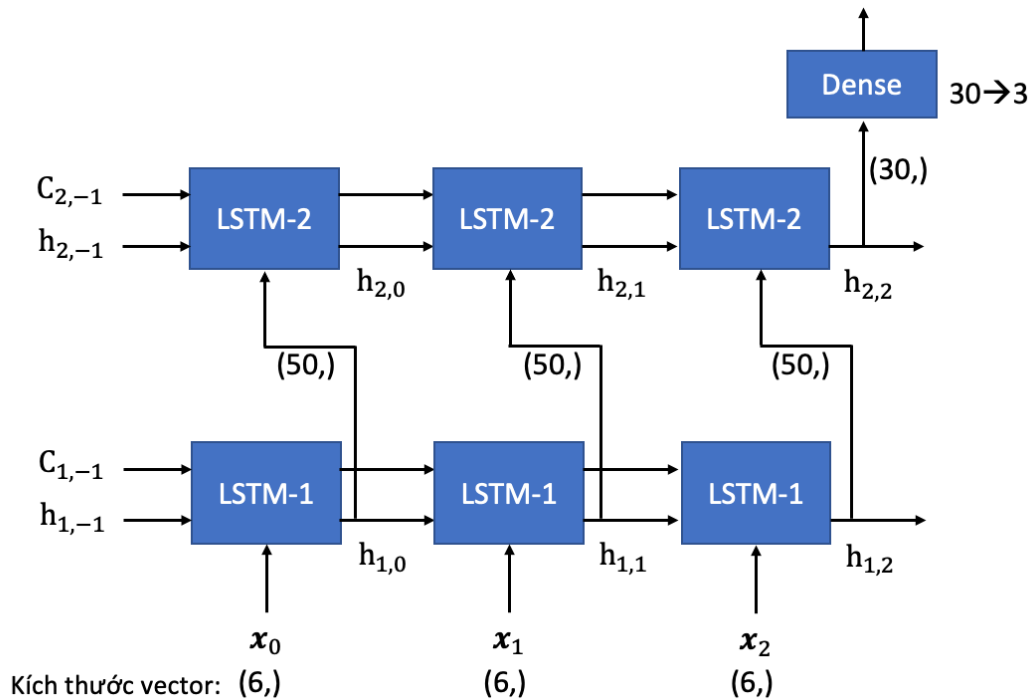
- **Pytorch:**

```
# Trong hàm khởi tạo của lớp định nghĩa mô hình Unet
import torch.nn as nn
self.conv_op3 = nn.Conv2d(64, K, kernel_size=1, stride=1, padding=0)
self.act_op3 = nn.Softmax(dim=1) # Pytorch shape: (N, C, H, W) => dim of channel = 1
```

```
# Trong hàm forward của lớp định nghĩa mô hình Unet
return self.act_op3(self.conv_op3(X)) # X là đầu ra của Op2
```

Câu 4. (L.O.1,2,3) (2 điểm): Cho đoạn mã sau trong Python

- a) **[0,5đ]** Vẽ sơ đồ của mô hình được định nghĩa trong đoạn mã sau đây. Hướng dẫn: sơ đồ ở dạng “unroll”, thể hiện từng time-step.



- b) **[0,5đ]** Xác định hình dạng dữ liệu của biến “output”
Shape: (4,3)

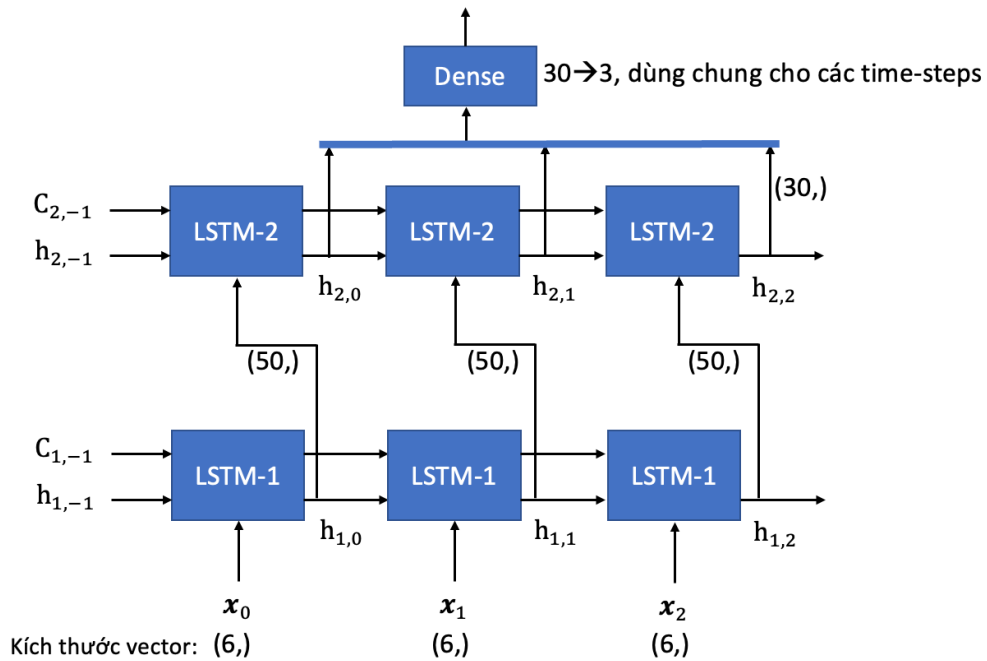
```
# Import libs
import numpy as np
from keras.models import Sequential
import keras.layers as layers
# Define model
model = Sequential([
    layers.LSTM(50, return_sequences=True),
    layers.LSTM(30, return_sequences=False),
    layers.Dense(3, activation="softmax")
])
model.compile(optimizer='adam', loss='crossentropy')
# Load (generate) batch data
N, T, F = 4, 3, 6
batch_data = np.arange(0, N * T * F).astype(np.float32).reshape(N, T, F)
# Do prediction
output = model.predict(batch_data)
```

- c) **[0,5đ]** Xác định hình dạng dữ liệu của biến “output” nếu thay `layers.LSTM(30, return_sequences=False)` bởi `layers.LSTM(30, return_sequences=True)`
Shape: (4,3,3)
- d) **[0,5đ]** Vẽ lại mô hình, nếu thay phần định nghĩa mô hình như đoạn mã sau đây:

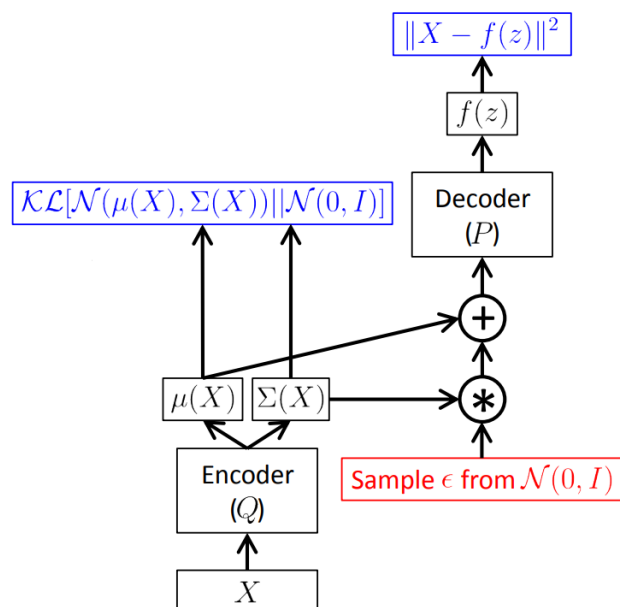

```

model = Sequential([
    layers.LSTM(50, return_sequences=True),
    layers.LSTM(30, return_sequences=True),
    layers.TimeDistributed(layers.Dense(3, activation="softmax"))
])

```



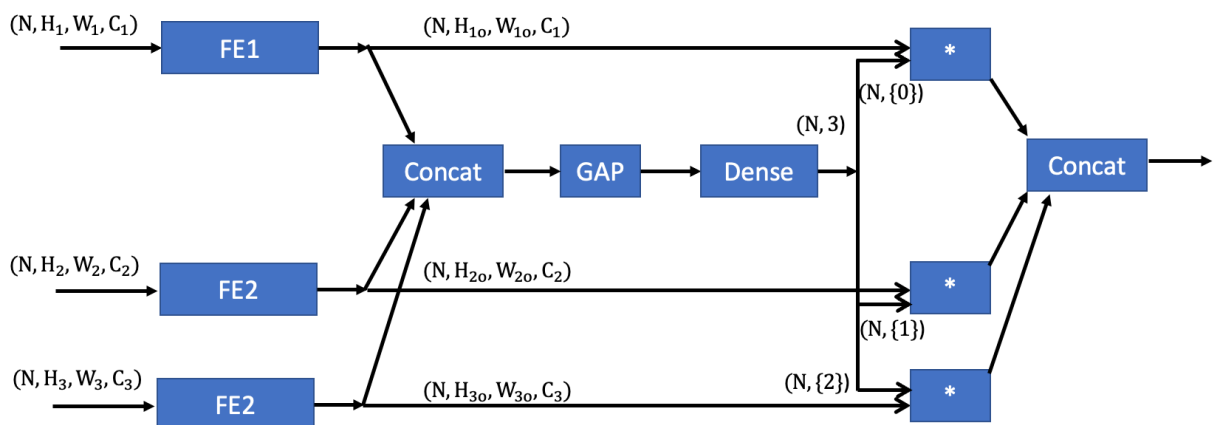
Câu 5. (L.O. 2,3) (1 điểm):



Hình 4. Mô hình VAE trong huấn luyện

- a) **[0,5đ]** Hãy giải thích ngắn gọn (vài dòng) về đầu ra của bộ Encoder và Decoder. Nêu được: chúng là gì? Ý nghĩa trong ứng dụng? Trong lúc huấn luyện bằng cách nào mà ép đầu ra về cái chúng ta mong muốn?
- **Đầu ra của Encoder:**
 - Encoder cố gắng học phân phối dữ liệu của tập huấn luyện. Nó giả sử rằng, phân phối có dạng phân phối chuẩn tắc đa biến (việc này luôn có thể, vì tồn tại một hàm toán để biến đổi giữa hai phân phối xác suất); do đó, đầu ra là: mean và covariance của dữ liệu trong tập huấn luyện.
 - **Đầu ra của Decoder:**
 - Khi đã có phân phối xác suất đã học được bởi Encoder, Decoder lấy mẫu phân phối này và biến đổi ngược ra miền ảnh.
 - Trong giai đoạn huấn luyện:
 - Nhờ sử dụng hàm mất mát là KL-Divergence mà chúng ta có thể ép đầu ra của Encoder về phân phối chuẩn tắc đa biến
 - Nhờ sử dụng hàm mất mát L2 giữa ảnh được sinh ra và ảnh gốc mà chúng ta có thể ép Decoder có thể biến đổi từ vector lấy mẫu ra ảnh.
- b) **[0,25đ]** Sau khi huấn luyện, bằng cách nào sử dụng?
- Sử dụng bộ Decoder để sinh ảnh, ảnh được sinh ra không nằm trong tập huấn luyện.
 - Bằng cách:
 - Lấy mẫu phân phối chuẩn tắc đa biến và đưa vào bộ Decoder
- c) **[0,25đ]** Hãy giải thích ngắn gọn (vài dòng): bằng cách nào sai số của bộ Decoder được sử dụng để điều chỉnh bộ Encoder?
- Bước Backward của giải thuật huấn luyện (SGD) lan truyền ngược sai số xuyên qua các phép toán * và + để vào bộ Encoder. Kết hợp với sai số của bộ Encoder (KL-Divergence), bước Backward xác định được lượng cần thay đổi trên tham số của Encoder để tiến hành bước cập nhật.

Câu 6. (L.O.2,3) (1 điểm): Cho đồ thị toán của một mô-đun như Hình 5.



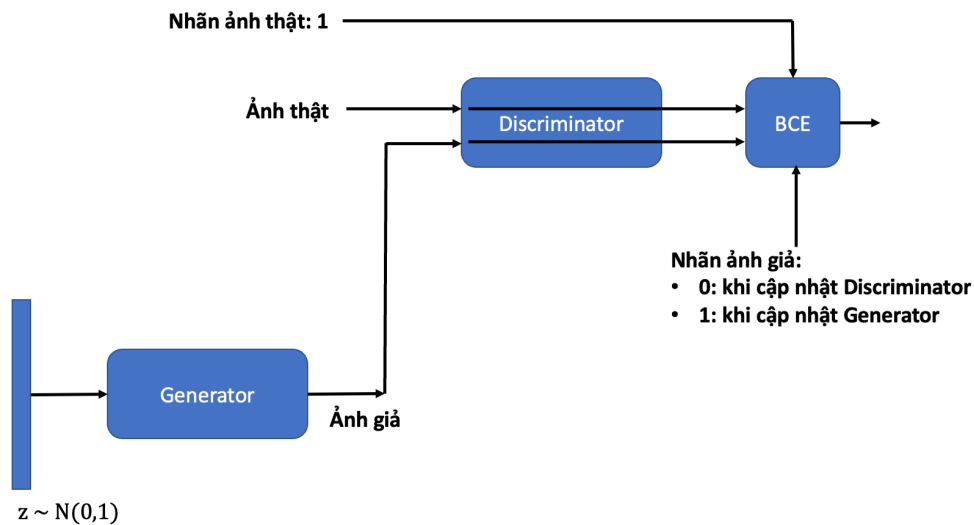
Hình 5. Đồ thị tính toán

Giả sử:

- $FE1(.)$, $FE2(.)$ là mô-đun rút trích đặc trưng đã được định nghĩa sẵn; có thể gọi hàm $FE1(.)$, $FE2(.)$
- Concat: nối theo chiều channel

- GAP: Global Average Pooling
Viết code (keras hoặc pytorch) định nghĩa mô-đun tính toán trong hình. Định nghĩa một hàm có 3 đầu vào, tính toán và trả về kết quả.

Câu 7. (L.O.2,3) (1 điểm): Cho mô hình GAN như Hình 6



Hình 6. Mô hình GAN

- a) **[0,5đ]** Nêu cách sử dụng mô hình trong Hình 6, sau khi đã huấn luyện.
- Sau khi huấn luyện, chúng ta có thể dùng Generator để sinh hình ảnh (giống ảnh thật), bằng cách:
 - Lấy mẫu vector z từ phân phối chuẩn tắc đa biến
 - Bộ Generator nhận z và ánh xạ sang miền ảnh.
- b) **[0,5đ]** Hãy viết ngắn gọn để giải thích lý do của cách huấn luyện sau đây: Khi cập nhật các tham số của khối “Discriminator” thì nhãn của ảnh thật là 1 và nhãn của ảnh giả là 0. Tuy nhiên, khi cập nhật các tham số của khối “Generator” thì: (a) khối “Discriminator” được đóng băng và (b) nhãn của ảnh giả phải là 1? Lưu ý: {1: thật; 0: giả}
- **Discriminator:** đóng vai trò là chuyên gia phân biệt thật vs giả. Do đó, để huấn luyện nó thì cần phân biệt ảnh thật (nhãn 1) và ảnh giả (nhãn 0) một cách rõ ràng trên tập ảnh huấn luyện; khi Generator làm việc tốt thì ảnh giả giống như thật và lúc đó Discriminator cần phải học thật tốt để phân biệt chúng.
 - **Generator:** đóng vai trò là chuyên gia làm giả (ảnh) để đánh lừa được chuyên gia Discriminator. Do đó, để huấn luyện Generator thì khi nó sinh được ảnh giả, thì cần báo (lừa) Discriminator rằng đấy là ảnh thật (nhãn 1)! và chúng ta cũng cần đóng băng bộ Discriminator khi huấn luyện Generator để Discriminator không học sai ảnh giả nhưng nhãn là ảnh thật (nhãn 1).

--- HẾT---