1. (10 points) **Haar-like Feature**

Haar-like features are popularly used for object detection, as mentioned in the class and references can be found in

https://en.wikipedia.org/wiki/Haar-like_feature
https://www.merl.com/publications/docs/TR2004-043.pdf

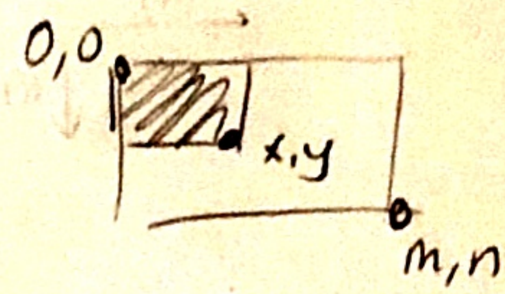
It can be efficiently implemented using integral image, so that the filter response at each location is at constant computation cost.

a) (5 points) Given an input image $I \in R^{m \times n}$, describe how to calculate the integral image, denoted by $II \in R^{m \times n}$.

The integral image at location x, y contains the sum of pixels above and to the left of it ie.

$$II(x,y) = \sum_{x' < x,\, y' < y} I(x', y'), \text{ repeat}$$

for all values in m×n



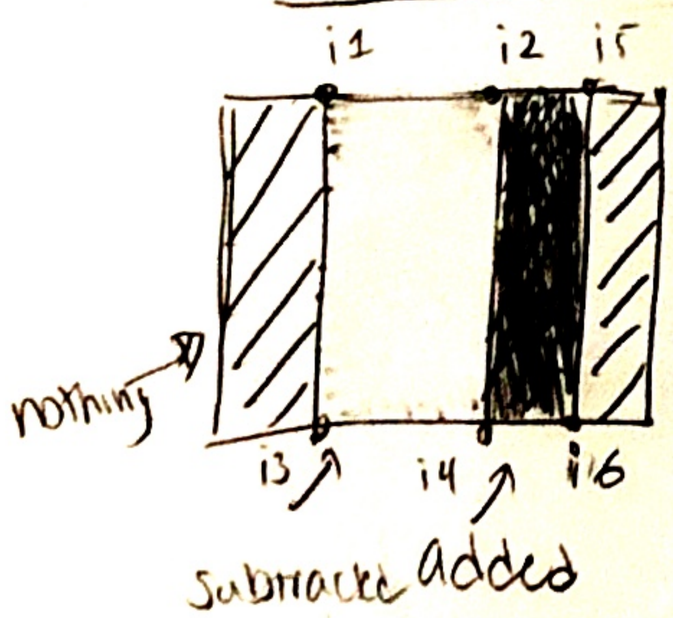
b) (5 points) For a Haar filter $H \in R^{3 \times 5}$ defined as below:

filter row

$$H = \begin{pmatrix} 0 & -1 & -1 & 1 & 0 \\ 0 & -1 & -1 & 1 & 0 \\ 0 & -1 & -1 & 1 & 0 \end{pmatrix}$$

rows, column

What is the equation to calculate the Haar features using integral image II?



Within the detection window, we would subtract the areas for the negatives and add the area for the positive

$$\sum_{j=0}^{5} \sum_{i=0}^{3} H[i,j]\, I[i,j]$$

or w/ the integral image, this would become:

$$(-II_4 - II_1 + II_2 + II_3) + (II_6 + II_2 - II_5 - II_4)$$

For negative area — For pos. area

1b cont.

If we say that the filter is applied to the top left corner of the image at (0,0), then this can be more specifically written as:

$$(-II[2,2] - II[0,0] + II[0,2] + II[2,1])$$
$$+ (II[2,3] + II[0,2] - II[0,3] - II[2,2])$$

Ideally I would only take into account the corner pieces like mentioned previously, but in order to accomodate restrictions due to image percision, I took the leftmost point

	0	1	2	3	4
0	0	-1	-1	1	0
1	0	-1	-1	1	0
2	0	-1	-1	1	0

2. (10 points) **Object Detection**

Object Detection is a basic visual task, trying to locate object(s) in an input image. Object Detection Methods can be classified into two categories, One stage methods and Two-stage methods.

a) (5 points) Explain the difference between them and name two representative methods for each category.

→ examples?

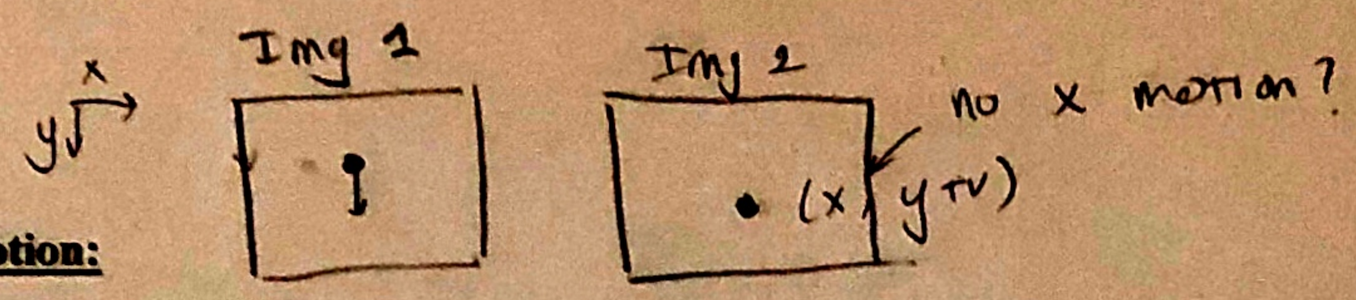
As indicated by the names, one stage object detection occurs in one stage while two stage occurs in two stages. Two stage has an additional stage for region proposal, while one stage has instead a region proposal net, which gets fed in directly to the rest of the network. One stage is end to end trainable. Two examples of double stage networks are R-CNN, and Fast R-CNN. Two examples of one stage networks are Faster R-CNN and Mask R-CNN.

b) (5 points) Explain the components of Faster-RCNN, how it improves over previous methods and how to train it step by step.

The components of faster-RCNN are a deep convolutional neural network, a region proposal network, region CNN features, and box offset regressor and softmax classifier.

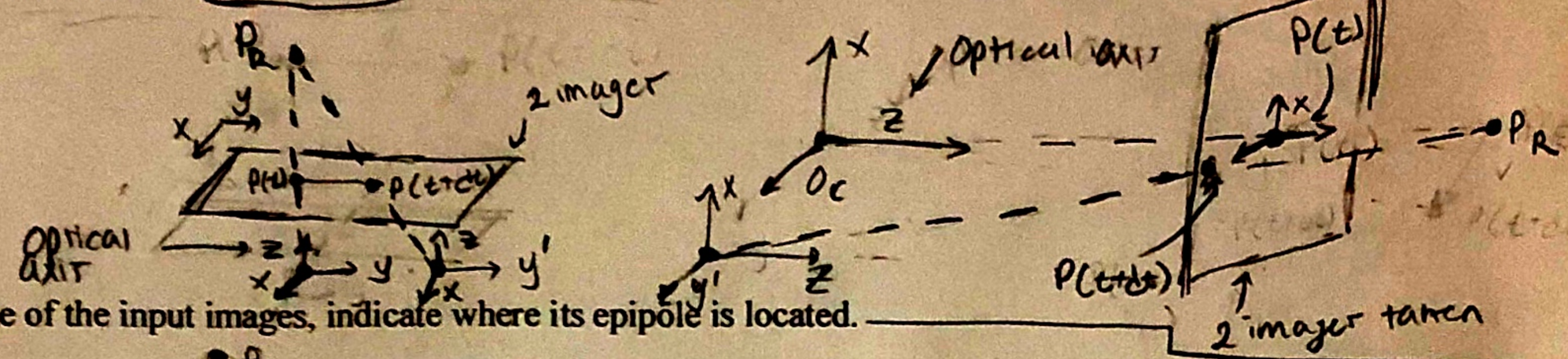
How it improves:
What a Region Proposal network does is it gets trained to propose region proposals directly. This makes the network is much faster and more accurate (compared to previous methods).

How to train step by step:
The way the network gets trained is by 1st taking an input image and passing it through the convolutional neural net. The CNN returns the feature maps for that image. After that the Region Proposal network (RPN) takes these features and produces object proposals. After, ROI pooling is done to get the features to be the same size. These proposals are then sent to the fully connected layer. What this does is it classifies (softmax) and also predicts the bounding boxes, with the box offset regressor predicting the offset.

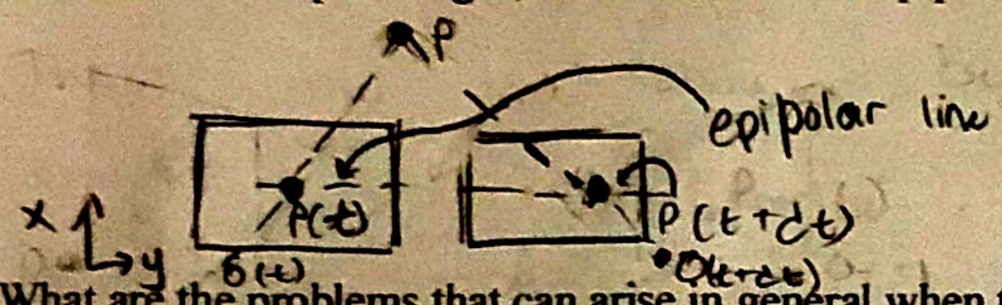


3. (24=4*6 points) **Motion:**

(a) Consider two images taken by a camera that translates down along the camera's y-axis direction while viewing a static scene. (Hence the camera's optical axis is perpendicular to the direction of motion.) What does the motion field look like for this motion?



(b) For one of the input images, indicate where its epipole is located.



The epipolar line is along the y axis since the camera only moves in one direction, the y direction.

(c) What are the problems that can arise in general when tracking edges? $\lambda_2 >> \lambda_1$

Displacement may have some ambiguity along the edge direction. This also means that the eigen values of the Lucas-Kanade problem have values $\lambda_2 >> \lambda_1$, therefore this suffers from the aperture problem.

(d) What are the problems that can arise in general when tracking a smooth patch?

but, this small difference can just simply be due to noise

It becomes hard to estimate accurately the displacement/motion. The eigen value solutions for the Lucas-Kanade problem will both be small. Additionally, this problem is more likeli to have issues arrising from noise. You can track a corner b/c it's distinguishable, especially from its surroundings. But for a smooth surface, you may think a part moved a certain direction

(e) What is the brightness constancy constraint?

Since time doesn't change that much and movement is relatively small, we can think of reflectance properties as static and the surface can locally be thought of as planar.

(f) Derive an expression for brightness constancy.

$$I(x, y, t-1) = I(x + u(x,y), y + v(x,y), t),$$

$$I(x, y, t-1) \approx I(x, y, t) + I_x \cdot u(x,y) + I_y \cdot v(x,y)$$

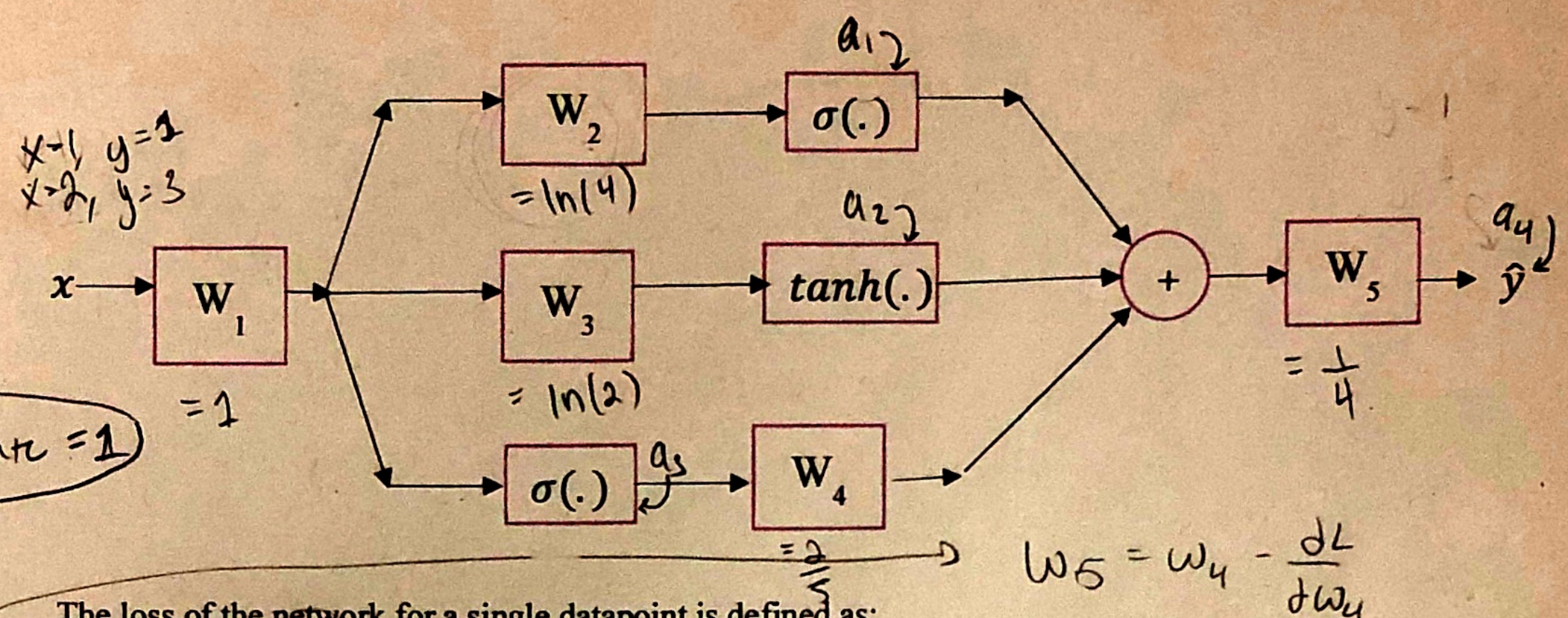
from Taylor expansion

$$I_x \cdot u + I_y \cdot v + I_t \approx 0$$

Since no displacement happens in the x direction, we can say

$$I_y \cdot v + I_t = 0$$

4. (20=4*5 points) **Backpropagation**
Consider the following network:



Learing rate =1

$w_5 = w_4 - \frac{\partial L}{\partial w_4}$

The loss of the network for a single datapoint is defined as:

$$L(y,\hat{y}) = \frac{(y-\hat{y})^2}{2}$$

Assume we are performing SGD with batch size 2 (average L) and the parameter update rule is given by $w_j^{n+1} = w_j^n - \frac{\partial L}{\partial w_j}$. For your current epoch, the two pairs of (input, output) samples are $(x = 1, y = 1)$ and $(x = 2, y = 3)$, the values of parameters are the following:

training input/output

x=1 y=1
x=2 y=3

$W_1 = 1, W_2 = \ln(4)\ , W_3 = \ln(2)\ , W_4 = \frac{2}{5}, W_5 = \frac{1}{4}$

$\frac{\partial L}{\partial \hat{y}} = \frac{1}{2}(2)(y-\hat{y})(-1) = \hat{y} - y$

Write the updated value of:

(i) W_1

$w_1 = w_1 - \frac{\partial L}{\partial w_1} = w_1 - \left(\frac{\partial L}{\partial \hat{y}}\right)\left(\frac{\partial \hat{y}}{\partial a_3}\right)\left(\frac{\partial a_3}{\partial w_1}\right)$

(ii) W_2

$w_2 = w_2 - \frac{\partial L}{\partial w_2} = w_2 - \left(\frac{\partial L}{\partial \hat{y}}\right)\left(\frac{\partial \hat{y}}{\partial a_1}\right)\left(\frac{\partial a_1}{\partial w_2}\right)$

(iii) W_3

$w_3 = w_3 - \frac{\partial L}{\partial w_3} = w_3 - \frac{\partial L}{\partial \hat{y}}\frac{\partial \hat{y}}{\partial a_2}\frac{\partial a_2}{\partial w_3} =$

(iv) W_4

$w_4 = w_4 - \frac{\partial L}{\partial w_4} = w_4 - \left(\frac{\partial L}{\partial \hat{y}}\right)\left(\frac{\partial \hat{y}}{\partial w_4}\right) = w_4 - \frac{\sum_i (\hat{y}-y)(w_5 a_3)}{2}$

(v) W_5

$w_5 = w_5 - \frac{\partial L}{\partial w_5} = w_5 - \left(\frac{\partial L}{\partial \hat{y}}\right)\left(\frac{\partial \hat{y}}{\partial w_5}\right) = w_5 - \frac{\sum (\hat{y}-y)(a_1 + a_2 + (w_4)(a_3))}{2}$

continued on sheet labeled 4.2

For forward prop. →

$a_1 = \sigma((w_2 + w_1)x)$

$a_2 = \tanh((w_3 + w_1)x)$

$a_3 = \sigma(w_1 x)$

$\hat{y} = a_4 = w_5(a_1 + a_2 + w_4 a_3)$

Continued on next sheet

4) cont.

$a_1 = \sigma((w_1 + w_2)x)$

When $x=1$: $a_1 = \sigma((\ln(4)+1)) = .916$

$x=2$: $a_1 = \sigma((\ln(4)+1)2) = .99$

$a_2 = \tanh((w_3 + w_4)x)$

$x=1$: $a_2 = \tanh((\ln(2)+1)) = .9345$

$x=2$: $a_2 = \tanh((\ln(2)+1)2) = .997713$

$a_3 = \sigma(w_1 x)$

$x=1$: $a_3 = \sigma(1) = .73105$

$x=2$: $a_3 = \sigma(2) = .88079$

$\hat{y} = a_4 = w_5(a_1 + a_2 + w_4 a_3)$

$x=1$: $a_4 = \frac{1}{4}(.916 + .9345 + (\frac{2}{5})(.73105)) = .53573 = \hat{y}_1$

$x=2$: $a_4 = \frac{1}{4}(.99 + .997713 + \frac{2}{5}(.88079)) = .58500725 = \hat{y}_2$

Please note, these are the approximations used later as well. for later calculations.

4.2

i) $w_1 = w_1 - \left(\frac{\partial L}{\partial \hat{y}}\right)\left(\frac{\partial \hat{y}}{\partial a_3}\right)\left(\frac{\partial a_3}{\partial w_1}\right) = w_1 - \frac{\sum_{i=1}^{2}(\hat{y}_i - y_i)}{2}(w_5 \cdot w_4)\,\sigma(w_1 x_i)(1-\sigma(w_1 x_i))\,x_i$

$= (1-\sigma(w_1 x))\,x_i = 1 - \frac{1}{2}[[.53573-1][\frac{1}{4}][\frac{2}{5}]\sigma[1]$

$(1-\sigma(1)) + [.58500725-3][\frac{1}{4}][\frac{2}{5}]\sigma[2][1-\sigma(2)]2]$

$\approx \boxed{1.03}$

ii) $w_2 = w_2 - \frac{\sum_{i=1}^{2}(\hat{y}_i - y_i)}{2}(w_5)\,\sigma((w_1+w_2)x_i)(1-\sigma((w_1+w_2)x_i))\,x_i$

$= \ln(4) - \frac{1}{2}[[.53573-1][\frac{1}{4}]\sigma((1+\ln(4)))[1-\sigma(1+\ln(4))] + [.58500725-3]$

$[\frac{1}{4}]\sigma[(1+\ln(4))2][1-\sigma((1+\ln(4))2)]2] \approx \boxed{1.40}$

iii) $w_3 = w_3 - \frac{\partial L}{\partial \hat{y}}\frac{\partial \hat{y}}{\partial a_2}\frac{\partial a_2}{\partial w_3} = w_3 - \frac{\sum_{i=1}^{2}(\hat{y}_i - y_i)}{2}(w_5)(1-\tanh(x_i(w_1+w_2))^2)x$

$= \ln(2) - \frac{1}{2}[[.53573-1][\frac{1}{4}][1-\tanh(1(1+\ln(4))^2)] + [.58500725-3]$

$[\frac{1}{4}][1-\tanh(2(1+\ln(4))^2)]2] \approx \boxed{.393}$

iv) $w_4 = w_4 - \frac{\sum_{i=1}^{2}(\hat{y}_i - y_i)}{2}(w_5\, a_{3i}) = \frac{2}{5} - \frac{1}{2}[[.53573-1][\frac{1}{4}][.73105] +$

$[.58500725-3][\frac{1}{4}][.88079]] \approx \boxed{.71}$

v) $w_5 = w_5 - \frac{\sum_{i=1}^{2}(\hat{y}_i - y_i)}{2}(a_{1i} + a_{2i} + (w_4)(a_{3i}))$

$= \frac{1}{4} - \frac{1}{2}([.53573-1](.916 + .9345 + (\frac{2}{5})(.73105)) +$

$[.58500725-3][.99 + .997713 + (.88079)(\frac{2}{5})])$

$\approx \boxed{3.57}$

Question 5i)

if we expand $y_\tau = \sum_{t=0}^{\tau} (-2)^t x_{\tau-1}$

we get ② $y_n = (-2)^0 x_n + (-2)^1 x_{n-1} + (-2)^2 x_{n-2} + \dots + (-2)^n x_0$

if we look at the other formula we are provided w/ (i.e. $y_t = w_1 x_t + w_2 y_{t-1}$):

$y_0 = w_1 x_0$

$y_1 = w_1 x_1 + w_2 y_0 = w_1 x_1 + w_2 (w_1 x_0)$

$y_2 = w_1 x_2 + w_2 y_1 = w_1 x_2 + w_2 (w_1 x_1 + w_2 w_1 x_0)$

$= w_1 x_2 + w_1 w_2 x_1 + w_2^2 w_1 x_0$

$\vdots$

① $y_n = w_1 x_n + w_2 y_{n-1} = w_1 x_n + w_2 w_1 x_{n-1} + w_2^2 w_1 x_{n-2} + \dots + w_2^n w_1 x_0$

$\therefore$ we can see that for ① and ② to have same values surrounding the x variables $\boxed{w_1 = 1 \text{ and } w_2 = -2}$

Ex. $w_1 x_n = -2^0 x_n \Rightarrow \therefore w_1 = 1$

$w_2 w_1 x_{n-1} = (-2)^1 x_{n-1} \Rightarrow \therefore$ since $w_1 = 1$, w_2 must $= -2$

Question 5ii) We can follow a similar procedure as i)

expanding y_τ, we get $y_n = a^0 x_n + a^1 x_{n-1} + a^2 x_{n-2} + \dots + a^n x_0$

$\therefore a^0 x_n = w_1 x_n \Rightarrow \therefore w_1 = 1$

$a^1 x_{n-1} = w_1 w_2 x_{n-1} \Rightarrow$ and since $w_1 = 1 \Rightarrow w_2 = a$

$\boxed{w_1 = 1 \; ; \; w_2 = a}$

Question Sili

Given: $y_t = f(z_t)$

$$= f(\overbrace{w_1 \ln x_t + w_2 z_{t-1}}^{z_t})$$

① $y_\tau = \prod_{t=0}^{\tau} x_{\tau-t}^{2^t}$

let $w_1 = 1$ and $w_2 = a \Rightarrow$ ② $y_t = f(\ln(x_t) + a z_{t-1})$

lets look at eqn ①:

$$y_\tau = \prod_{t=0}^{\tau} x_{\tau-t}^{2^t} \Rightarrow y_1 = \prod_{t=0}^{1} x_{\tau-t}^{2^t} = (x_1^{2^0})(x_0^{2^1})$$

$$y_n = \prod_{t=0}^{n} x_{\tau-t}^{2^t} = (x_n^{2^0})(x_{n-1}^{2^1})(x_{n-2}^{2^2}) \cdots (x_1^{2^{n-1}})(x_0^{2^n})$$

looking at eqn ②:

$$y_t = f(\ln(x_t) + a z_{t-1})$$

$$y_0 = f(\ln(x_0))$$

$$y_1 = f(\ln(x_1) + a z_0) = f(\ln(x_1) + a[\ln(x_0)])$$

$$y_2 = f(\ln(x_2) + a z_1) = f(\ln(x_2) + a(\ln(x_1) + a z_0))$$

$$= f(\ln(x_2) + a(\ln(x_1) + a(\ln x_0)))$$

$$= f(\ln(x_2) + a\ln(x_1) + a^2 \ln(x_0))$$

$\vdots$

$$y_n = f(\ln(x_n) + a\ln(x_{n-1}) + a^2 \ln(x_{n-2}) + \dots + a^n \ln(x_0))$$

if we look @ $y=1$:

$$(x_1^{2^0})(x_0^{2^1}) = f(\ln(x_1) + a\ln(x_0)) = f(\ln(x_0^a x_1))$$

$x^0 \quad x_0^2$

$\hookrightarrow$ for $y_2 = f(\ln(x_2 x_1^a x_0^{a^2})) \Rightarrow y_2 = x_2^{2^0} x_1^{2^1} x_0^{2^2}$

$$f\left(\sum_{i=0}^{n} \ln(x_{n-i}) a^i\right) = \prod_{i=0}^{n} x_i^{2^{n-i}}$$

$\boxed{f \text{ can be } e(\cdot) \text{ with the constraint that } a = 2}$

$\exp(\cdot) \Rightarrow$ $\boxed{\text{ie. } f(x) = \exp(x) \text{ st. } a=2}$