1.)
a) Given $I \in R^{m \times n}$

let's say $I = \begin{bmatrix} a & b \\ c & d \end{bmatrix}_{2\times 2}$

then integral image is calculated as follows:
Value at any point (x, y) is the sum of all pixels above up-left region to it.

So, $\underline{II} = \begin{bmatrix} a & b+a \\ c+a & a+b+c+d \end{bmatrix}$

$$I(x,y) = \sum_{\substack{x' \le x \\ y' \le y}} i(x', y')$$ where $i(x,y)$ is pixel value at (x,y)

b.) In given filter.

black region.

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

white region

Haar feature is present only when $\Delta >$ threshold.

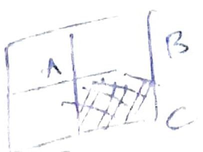
Δ = mean(black region) - Mean(white region).

$\underline{II}$ is integral image.

Sum of black region : $\underline{II}(3,4) - \underline{II}(3,3)$

Sum of white region : $\underline{II}(3,3) - \underline{II}(3,1)$

total sum = $\underline{II}(3,4) - \underline{II}(3,3) + \underline{II}(3,3) - \underline{II}(3,1) = \underline{II}(3,4) - \underline{II}(3,1)$



sum of shaded region $\Rightarrow$ $\underline{II}(C) + \underline{II}(A) - \underline{II}(B) - \underline{II}(D)$

2.) a.) One stage vs Two-stage.

One stage	Two-stage
i) They treat object detection as simple regression problem by taking an input image and learning class probabilities and bounding box co-ordinates	(i) Use a Region Proposal Network to generate regions of interests in first stage
	(ii) Send the region proposals down the pipeline for object classification and bounding box regression.
Low accuracy rates but faster	High accuracy rate but typically slower.
Ex: YOLO (You Only Look Once), SSD (Single Shot Multibox Detector).	Ex: Faster R-CNN, Mask RCNN, R-FCN

b.) Faster R-CNN → 3 parts

i) Convolution Layers: We train filters to extract appropriate features from the image.

ii) Region Proposal Network (RPN): It is a small neural network sliding on the last feature map of convolution layers and predict whether there is an object or not and also predict bounding box of those objects.

III) Classes & Bounding boxes prediction: We use a fully connected neural network that takes as input regions proposed by RPN & predict object class (Classification) & Bounding boxes (Regression).

Main improvement: Use of RPN instead of selective search.

Test-time speed of faster RCNN (0.2 sec), fast RCNN (2.3 secs) & RCNN (49 sec

Faster RCNN can be used for real-time object detection.

Training faster RCNN:

i) Train RPN, initialized with ImageNet pretrained model.

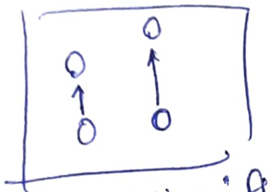
ii) Train a separate detection network by Fast RCNN using proposals generated by step 1 RPN.

iii) Fix convolutional layer, fine tune unique layers to RPN, initialized by detector network in step 2.

iv) fix conv. layers, fine tune fully connected layer of fast RCNN.

3.) a.) Translation in +y direction.

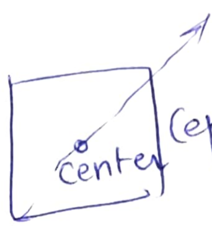
Motion flow field is in same direction.(~~up~~ vertical +y axis).



Motion flow field.

b) All planes include the line that connects the two focal points. Where this line intersects an image plane, that image point is epipole.

Since the depth does not change, the epipole should be located at center of input images.



(epipolar line crossing center)

c) When tracking edges, the edges cannot be determined unambiguously if it is viewed through a small aperture such that the ends of stimulus are not visible. This is known as aperture problem which can be solved using motion based prediction.

The other problems are non rigid object structures, occlusions & camera motion.

d.) When we have a smooth patch, gradients in that region have small magnitude (small λ_1, small λ_2) as in a low-texture region. It becomes difficult to track motion with smooth regions so we use high texture (edges) ~~but~~ which causes aperture problem.

e.) Brightness Constancy Constraint:
To match image brightness values across images $\Rightarrow$ very restrictive.
Assumption: intensity of surface patch remains constant.

f.) Brightness constancy expression:
Let a ~~pixel~~ (x,y) displaced (u,v). New pixel $(x+u, y+v)$

~~Brightness~~ Constancy Equation is

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

Using taylor expansion.

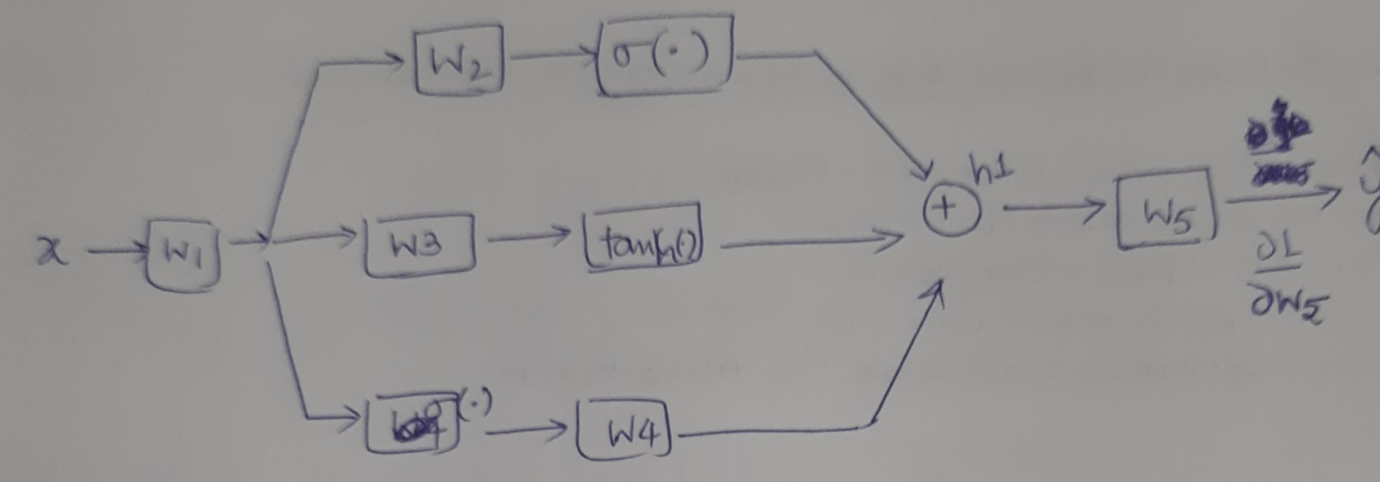
$$I(x,y,t-1) \approx I(x,y,t) + I_x \cdot u(x,y) + I_y \cdot v(x,y)$$

$I_x \leftarrow$ image derivate along x, $I_t \leftarrow$ difference over frames.

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

$$\boxed{\nabla I \cdot (u,v) + I_t = 0}$$

4)



$$\hat{y} = W_5\left(\sigma(W_2(W_1 x)) + \tanh(W_3(W_1 x)) + W_4\sigma(W_1 x)\right)$$

$W_1 = 1,\ W_2 = \ln(4).\ W_3 = \ln(2),\ W_4 = 2/5,\ W_5 = 1/4$

for $x=1$,

$$\hat{y} = 1/4\left(\sigma(\ln(4)) + \tanh(\ln(2)) + 2/5\ \sigma(1)\right)$$

$$\sigma(x) = \frac{1}{1+e^{-x}},\quad \tanh(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}.$$

$$\hat{y}(x=1) = 1/4\left(0.8 + 0.6 + 2/5\,\sigma(1)\right) = 0.42$$

for $x=2$

$$\hat{y} = 1/4\left(\sigma(2\ln(4)) + \tanh(2\ln(2)) + 2/5\,\sigma(2)\right)$$

$$\hat{y} = 1/4\left(0.9411 + 0.8823 + 2/5(0.88)\right)$$

$$\hat{y} = 0.543$$

Loss for this batch $= \frac{(1-0.42)^2}{2} + \frac{(3-0.543)^2}{2} = 3.1866$

$$\frac{\partial \hat{y}}{\partial W_5} = \sigma(W_2(W_1 x)) + \tanh(W_3(W_1 x)) + W_4\,\sigma(W_1 x))$$

for I_1, $$\frac{\partial \hat{y}}{\partial W_5} = \sigma(\ln(4)) + \tanh(\ln(2)) + 2/5\,\sigma(1)$$

$$= 0.8 + 0.6 + 0.2924 = 1.6924$$

$$\partial E/\partial W_5 = \frac{\partial E}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial W_5} = (-1)(y-\hat{y})\frac{\partial \hat{y}}{\partial W_5} = -0.9815$$

for I_2, $$\frac{\partial \hat{y}}{\partial W_5} = \sigma(2\ln(4)) + \tanh(2\ln(2)) + 2/5\,\sigma(2)$$

$$= 0.9412 + 0.8823 + 0.3522 = 2.175$$

$$\partial E/\partial W_5 = \cancel{1.2649}\ -5.343$$

$$\partial E/\partial W_5\,(\text{Avg}) = -3.162 \Rightarrow \boxed{W_5 = 1/4 + 3.162 = 3.412}$$

$\underline{W4}$: for I_1:

$$\frac{\partial E}{\partial W_4} = (y-\hat{y})(W_5\,\sigma(W_1 x)) = 1/4\,\sigma(1)\cdot(-0.58) = -0.106$$

for I_2

$$\frac{\partial E}{\partial W_4} = 1/4\,\sigma(2)(-2.457) = -0.5409$$

$$\frac{\partial E}{\partial W_4}(\text{Avg}) = -0.3234 \Rightarrow \boxed{W_4 = 2/5 + 0.3234 = 0.7234}$$

$\underline{W3}$

$$\frac{\partial E}{\partial W_3} = W_5 * (1 - \tanh^2(W_1 W_3 x)) \times W_1 x * (\hat{y} - y)$$

I_1: $\partial E/\partial W_3 = -0.0928$ $\quad$ I_2: $\partial E/\partial W_3 = -0.2721$

$$\partial E/\partial W_3\,(\text{Avg}) = -0.18248 \Rightarrow \boxed{W_3 = \ln(2) + 0.18248 = 0.87562}$$

$\underline{\underline{W_2}}$

$$\frac{\partial E}{\partial W_2} = (\hat{y}-y) W_5 \sigma(W_1 W_2 x)(1-\sigma(W_1 W_2 x)) W_1 x$$

I_1: $\partial E/\partial W_2 = -0.0232$, I_2: $\partial E/\partial W_2 = -0.06809$

$\partial E/\partial W_2 \text{ (Avg)} = -0.04564 \Rightarrow \boxed{W_2 = \ln(4) + 0.04564 = 1.43194}$

$\underline{\underline{W1}}$

$$\frac{\partial E}{\partial W_1} = (\hat{y}-y) W_5 (\sigma(W_1 W_2 x) \cdot (W_2 x)(1-\sigma(W_1 W_2 x)) + (1 - \tanh^2(W_1 W_3 x)) W_3 x$$
$$+ W_4 \sigma(W_1 x)(1-\sigma(W_1 x)) x)$$

I_1:

$$\frac{\partial E}{\partial W_1} = (-0.58)\left(\frac{1}{4}\right)\left[(0.8)(\ln(4))(0.2) + (0.64)(\ln(2)) + \frac{2}{5}(0.7310)(1-0.7310)\right]$$

$$= (-0.58)(1/4)[0.2218 + 0.44361 + 0.07865]$$

$$= -0.10788$$

I_2:

$$\partial E/\partial W_1 = -2.457\left(\frac{1}{4}\right)[(0.9411)(2\ln(4))(0.0589) + 0.22145 \times 2 \ln 2$$
$$+ \frac{2}{5}((0.8807)(0.1193)(2))$$

$$= -2.457\left(\frac{1}{4}\right)[0.15368 + 0.3069 + 0.8405]$$

$$= -0.3345$$

$\partial E/\partial W_1 \text{ (Avg)} = -0.22121 \Rightarrow \boxed{W_1 = 1 + 0.22121 = 1.22121}$

$\Rightarrow$ i) $W_1 = 1.22121$

ii) $W_2 = 1.43194$

iii) $W_3 = 0.87562$

iv) $W_4 = 0.7234$

v) $W_5 = 3.412$.

5.) $y_t = W_1 x_t + W_2 y_{t-1}$ — (1)

i) $y_\tau = \sum_{t=0}^{\tau} (-2)^t x_{\tau-t}$

Finding relation b/w y_τ & $y_{\tau-1}$

$y_\tau = x_\tau + (-2)^1 x_{\tau-1} + (-2)^2 x_{\tau-2} + \cdots + (-2)^{\tau-1} x_1 + (-2)^\tau x_{\tau-0}$

$\cancel{y_{\tau-1} = x_\tau + (-2)^1 x_{\tau-1} + \cdots + (-2)^{\tau-1} x_1}$

$\cancel{y_\tau - y_{\tau-1} = (-2)^\tau x_\tau}$

$y_{\tau-1} = (-2)^0 x_{\tau-1} + (-2)^1 x_{\tau-2} + (-2)^2 x_{\tau-3} + \cdots$

$y_\tau - (-2) y_{\tau-1} = x_\tau$

$y_\tau = x_\tau + -2 y_{\tau-1}$

∴ Substituting in (1) we get, $\boxed{W_1 = 1,\ W_2 = -2}$

ii) $y_\tau = \sum_{t=0}^{\tau} a^t x_{\tau-t}$, $a > 0$

$y_{\tau-1} = a^0 x_{\tau-1} + a^1 x_{\tau-2} + \cdots$

$y_\tau = a^0 x_\tau + a^1 x_{\tau-1} + a^2 x_{\tau-2} + \cdots$

$y_\tau - a y_{\tau-1} = a^0 x_\tau$

$y_\tau = x_\tau + a y_{\tau-1} \Rightarrow \boxed{W_1 = 1,\ W_2 = a}$

iii) $z_t = w_1 \ln x_t + w_2 z_{t-1}$

$w = 1, \ w_2 = a$

$y_t = f(z_t)$.

$y_\tau = \prod_{t=0}^{\tau} x_{\tau-t}^{2^t}$; $\ln(y_\tau) = \sum_{t=0}^{\tau} 2^t \ln x_{\tau-t}$

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

$\therefore \ln y_\tau - 2\ln(y_{\tau-1}) = \ln x_\tau$

$\ln y_\tau = \ln x_\tau + 2 \ln y_{\tau-1}$

$a \ln y_\tau = a \ln x_\tau + 2a \ln y_{\tau-1}$

$\ln y_\tau = \ln x_\tau + 2 \ln y_{\tau-1}$ $\Rightarrow$ Here $a = 2$. (from ii)

& $z_\tau = \ln y_\tau$

$y_\tau = e^{z_\tau} \Rightarrow f(\cdot) = e^{(\cdot)}$

$f(x) = e^x$.