CSE527 Introduction to Computer Vision Spring 2021

Final

Due 23:59pm EST, May 18, 2021

Name:
ID:

There are 100 points in the exam.
Use your time wisely. Most questions require brief answers.

1	
2	
3	
4	
5	
6	
Total	

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 $I \in \boldsymbol{R}^{m \times n}$, describe how to calculate the integral image, denoted by $II \in R^{m \times n}$.

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

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

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

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.

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

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.

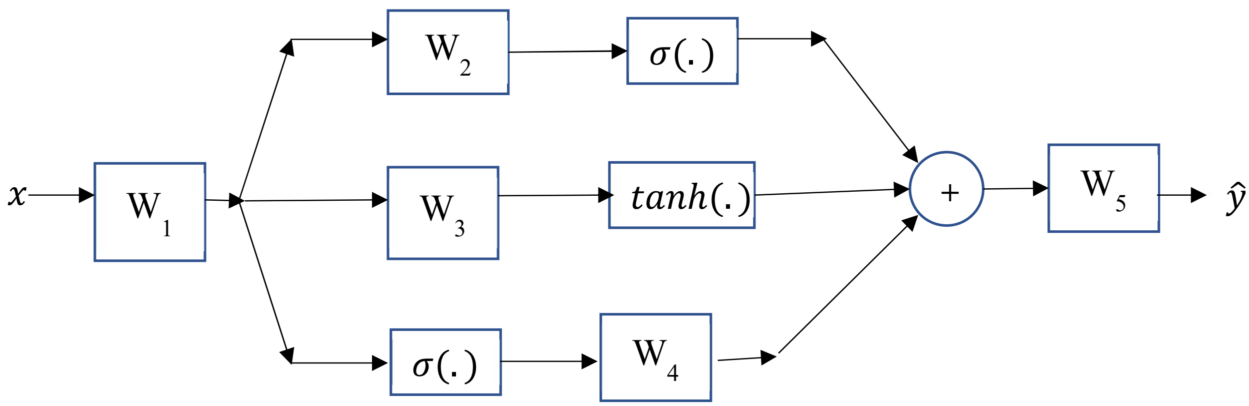
(c) What are the problems that can arise in general when tracking edges?

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

(e) What is the brightness constancy constraint?

(f) Derive an expression for brightness constancy.

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



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 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:

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

Write the updated value of:

$(i)\ W_1$

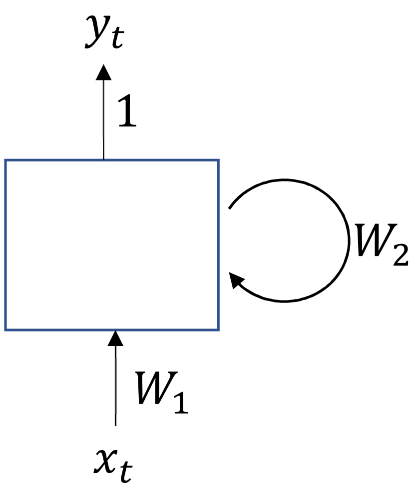
$(ii)\ W_2$

$(iii)\ W_3$

$(iv)\ W_4$

$(v)\ W_5$

5. (16=8*2 points) **Standard RNN:**



The input output of the following network is given by:

$$y_t = W_1 x_t + W_2 y_{t-1}$$

(i) If we want to implement the following input-output mapping:

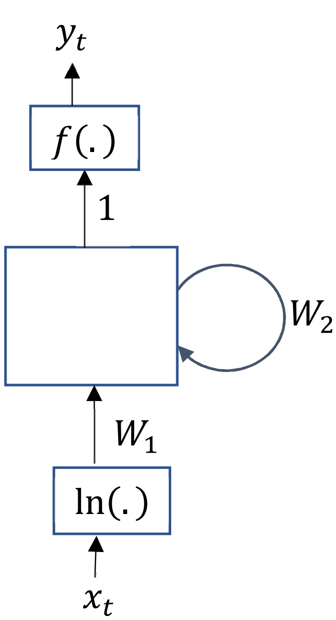
$$y_\tau = \sum_{t=0}^{\tau} (-2)^t x_{\tau-t}$$

What should be the value of W_1 and W_2 ?

(ii) If we want to implement the following input-output mapping:

$$y_\tau = \sum_{t=0}^{\tau} a^t x_{\tau-t} \quad \text{for some } a > 0$$

What should be the value of W_1 and W_2 ?



(iii) The input output relation of the network in the right is given by:

$$z_t = W_1 \ln x_t + W_2 z_{t-1}$$
$$y_t = f(z_t)$$

If we want to implement the following input-output mapping:

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

Where W_1 and W_2 are the weights obtained from part (ii), what should be the function $\boldsymbol{f(.)}$ to implement the above relation?

6. (20 points) **<u>Visual object tracking.</u>**
See the other file.