DSAA ASSIGNMENT-2

AAYUSH GOEL
20171188

REPORT

Problem 1 :-

a) In electronics and signal processing, a Gaussian filter is a filter whose impulse response is a Gaussian function.

To create the gauss filter of size N, the gaussian formula for 2D has been used where sigma represents the standard deviation.

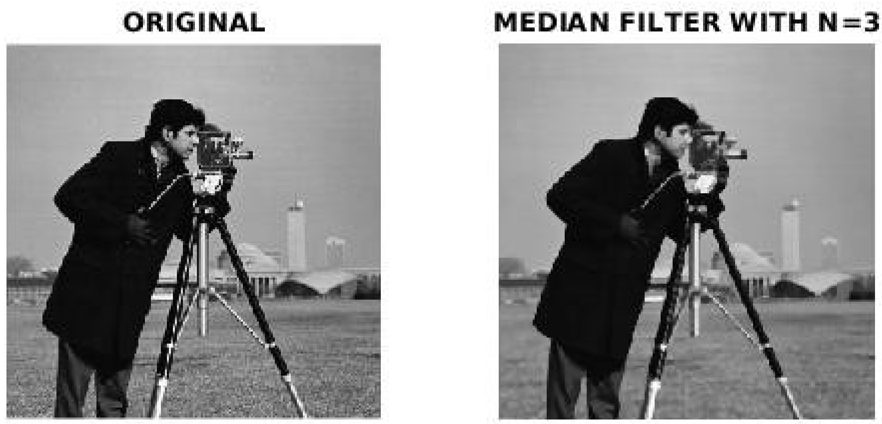
$$g(x, y) = \frac{1}{2\pi\sigma^2} \cdot e^{-\frac{x^2+y^2}{2\sigma^2}}$$

Gaussian Implemented vs fspecial at N = 64 and sigma = 2

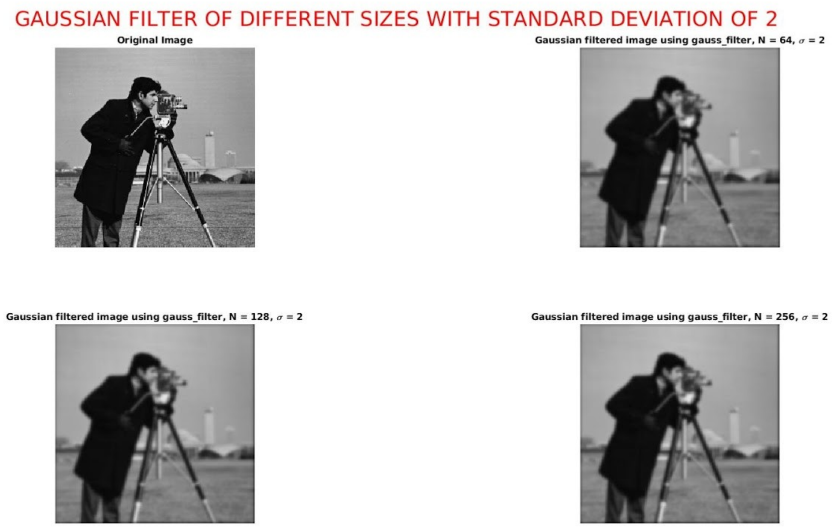
Gaussian Implemented vs fspecial at N = 32 and sigma = 2

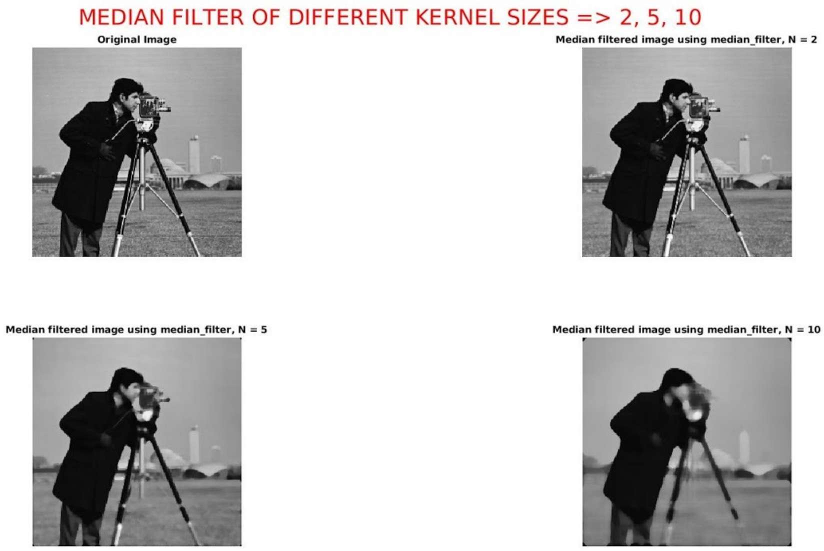
(b) The main idea of the median filter is to run through the signal padded with zeros, entry by entry, replacing each entry with the median of neighboring entries. The pattern of neighbors is called the "window", which slides, entry by entry, over the entire signal.

A window of size NxN has been slided over the entire image and the elements in the window have been sorted and the middle element is picked and then, the window is replaced by the median element.

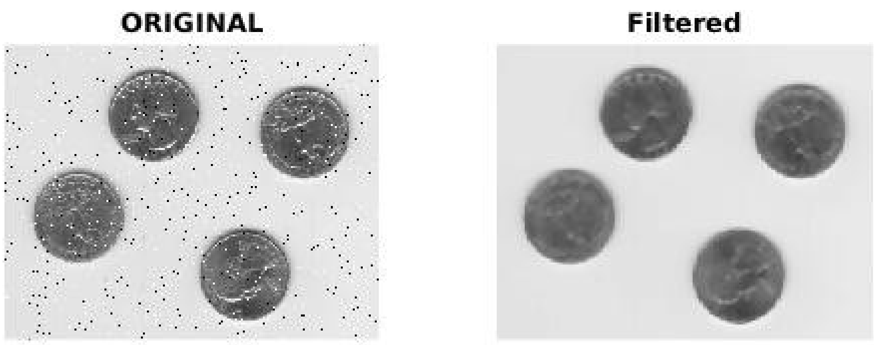


(c)



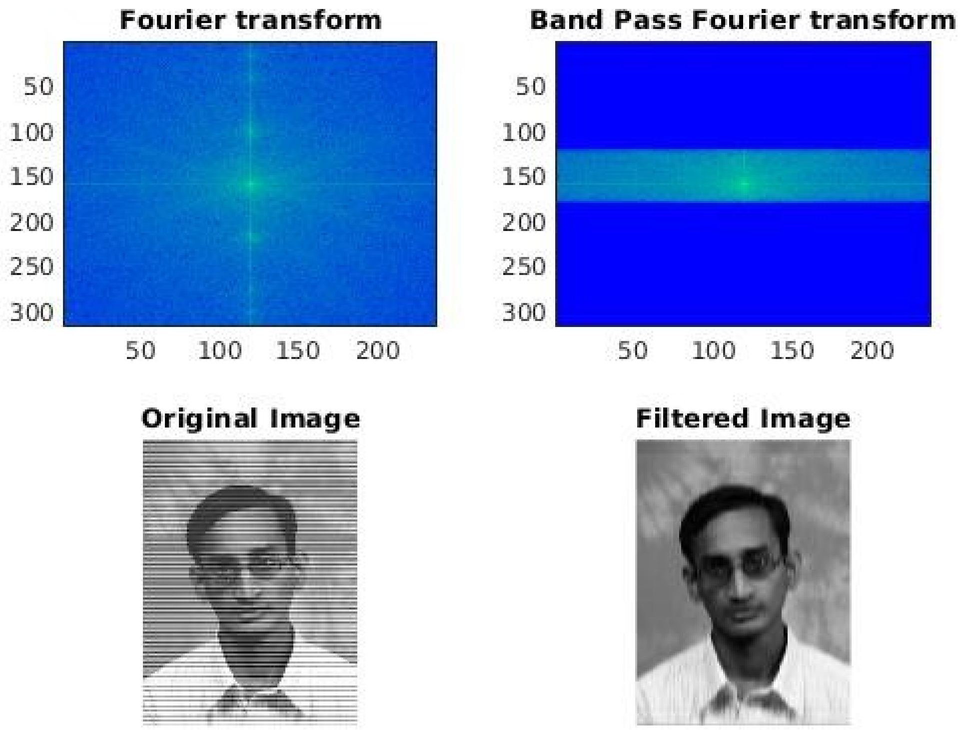


(d) The median filter is particularly effective for speckle noise and salt and pepper noise (impulsive noise that is noise which includes unwanted, almost instantaneous sharp sounds) because they appear for short duration and while applying the median we replace a small part containing the noise with the median value of the window.



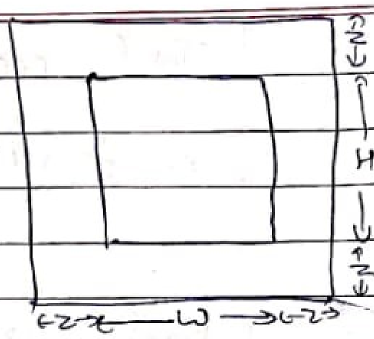
(e) In the given image, to remove the noise band-pass filters are applied

Band Pass Filter - Some of the high frequencies which contributed to the noise are not allowed to pass through the filter. This preserves the image but the noise is properly removed.



Problem 2 :-

Original width = W
After padding = W + 2z
Actual width that can be traversed due to filter = W + 2z − f



After one convolution

$$w(1) = \frac{w + 2z - f}{s} + 1$$

This will result in a recursion relation

$$w(i) = \frac{w(i-1) + 2z - f}{s} - 1$$

$$w(2) = \frac{w(1) + 2z - f}{s} + 1$$

$$w(2) = \frac{\frac{W + 2z - f}{s} + 1 + 2z - f + 1}{s}$$

$$= \frac{w + (2z - f + s)(1+s)}{s^2}$$

$$w(N) = \frac{w + (2z - f + s)(1 + s + s^2 + \dots + s^{N-1})}{s^N}$$

$$w(N) = \frac{w + (2z - f + s)\left(\frac{s^N - 1}{s-1}\right)}{s^N}$$

Illy, $H(N) = \frac{H + (2z - f + s)\left(\frac{s^N - 1}{s-1}\right)}{s^N}$

⇒ Channels will remain same since it's applied to all channels

Multiplication ⇒ F^2 at each step in convolution
Addition ⇒ $F^2 - 1$ at each step in a convolution

Total operations after N convolutions

$$\Rightarrow \sum_{i=1}^{N} H(i)w(i)\,[f^2+f^2-1] \times \text{channels}$$

Final size

$$\Rightarrow \left(\frac{w+(2z-f+s)\left(\frac{s^N-1}{s-1}\right)}{s^N},\ \frac{h+(2z-f+s)\left(\frac{s^N-1}{s-1}\right)}{s},\ \text{channels} \right)$$

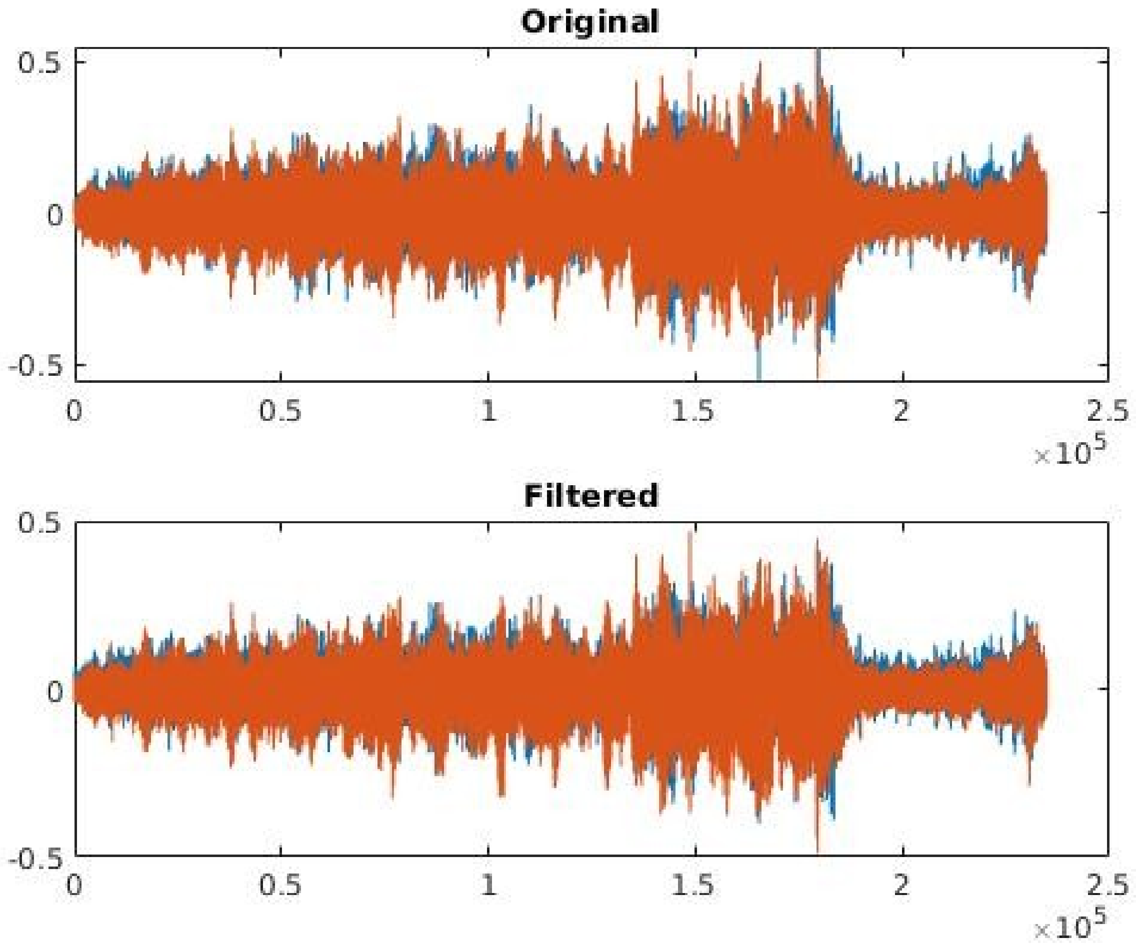
$$\text{Total operations} = \sum_{i=1}^{N} H(i)\,w(i)\,[2f^2-1] \times \text{channels}$$

$$= \sum_{i=1}^{N} \left(\frac{w+(2z-f+s)\left(\frac{s^i-1}{s-1}\right)}{s^i} \right) \left(\frac{h+(2z-f+s)\left(\frac{s^i-1}{s-1}\right)}{s^i} \right) \times [2f^2-1] \times \text{channels}$$

Problem 3 :-

We have divided the signal into multiple samples based on the duration on the smaller samples. We perform a correlation with each of the 10 possible tunes and the number corresponding to the max value of correlation is output as the solution and this has been carried out for all the samples.

Problem 4:-



A bandpass filter has been used centred around the zero as the noise will have low amplitudes in comparison to the original signal without the noise. The audio signal has high amplitudes which can be easily separated from low amplitude noise by using bandpass filter along with fft.

Problem 5 :-

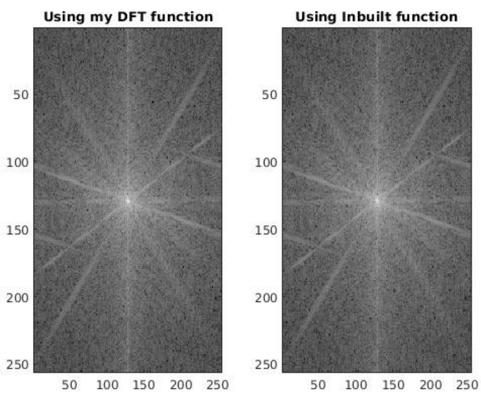
Size	My FFT	My DFT	In-built FFT
4096	18 sec	15 sec	0.9sec
2048	4.4sec	1.8sec	0.3sec
1024	1.05sec	0.3sec	0.08sec
cameraman	0.18sec	0.11sec	0.03sec

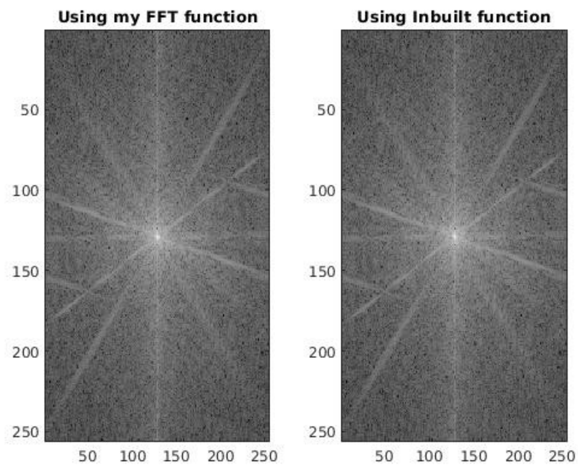
Pseudo Code

RECURSIVE-FFT(a)

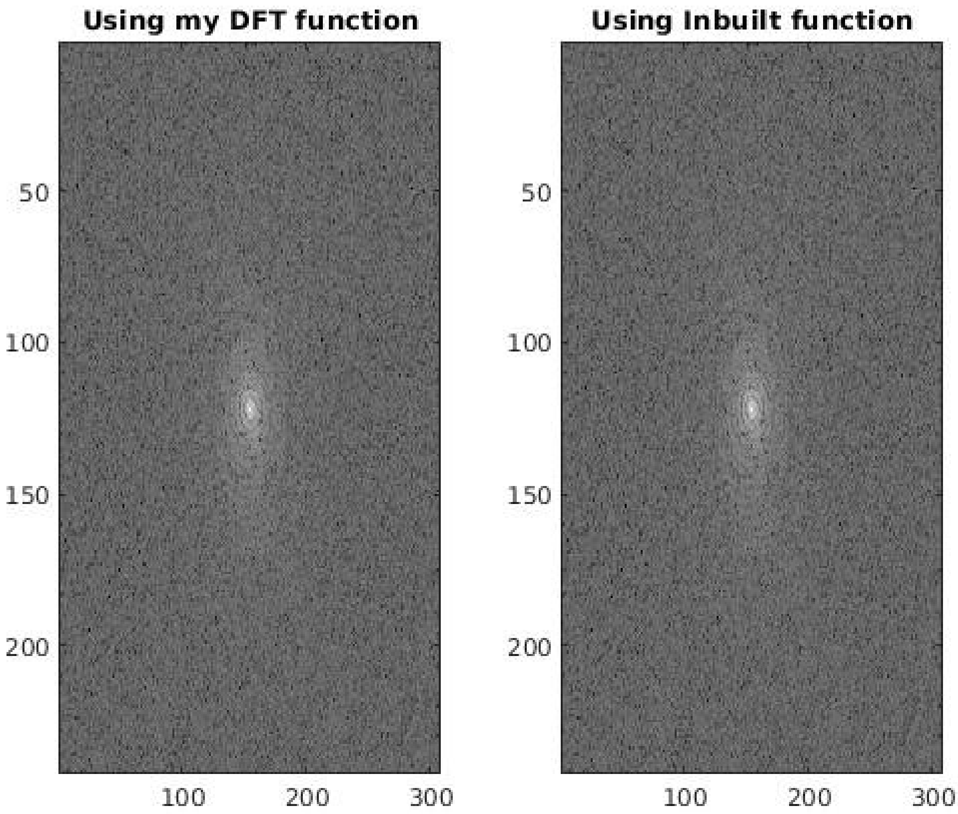
$n \leftarrow length[a]$ $\qquad \triangleright$ n is a power of 2.
if $n = 1$
    then return a
$\omega_n \leftarrow e^{2\pi i/n}$
$\omega \leftarrow 1$
$a^{[0]} \leftarrow (a_0, a_2, \ldots, a_{n-2})$
$a^{[1]} \leftarrow (a_1, a_3, \ldots, a_{n-1})$
$y^{[0]} \leftarrow$ RECURSIVE-FFT($a^{[0]}$)
$y^{[1]} \leftarrow$ RECURSIVE-FFT($a^{[1]}$)
for $k \leftarrow 0$ **to** $n/2 - 1$
    do $y_k \leftarrow y_k^{[0]} + \omega\, y_k^{[1]}$
        $y_{k+(n/2)} \leftarrow y_k^{[0]} - \omega\, y_k^{[1]}$
        $\omega \leftarrow \omega\, \omega_n$
return y $\qquad \triangleright$ y is assumed to be column vector.

cameraman.tif





pic.jpeg


Using my DFT function
Using Inbuilt function
50
100
150
200
100
200
300
50
100
150
200
100
200
300

Problem 6:-

(6) If W_n is the FFT Matrix

Applying double fft equals multiplying by W_n^2

$$W = W_n \times W_n = \frac{1}{\sqrt{N}}\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ \vdots & & & & & \\ 1 & & \cdots & & & \end{bmatrix} \frac{1}{\sqrt{N}}\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ & & e & & \\ 1 & & & & \end{bmatrix}$$

For any cell (i,j)

$$W_{ij} = \sum_{k=0}^{n} e^{-\frac{2\pi i}{N} ik} \cdot e^{-\frac{2\pi j}{N} kj}$$

$$= \sum_{k=0}^{n} e^{-\frac{2\pi i}{N} k(i+j)}$$

if $i+j = N$

$$W_{ij} = \frac{1}{N}\sum_{k=0}^{n} e^{-\frac{2\pi i k (N)}{N}} = 1$$

else

$$W_{ij} = \frac{1}{N}\sum_{k=0}^{n} e^{-\frac{2\pi i k}{N}(x)} \qquad \text{where } x = (i+j)$$

$$= 0$$

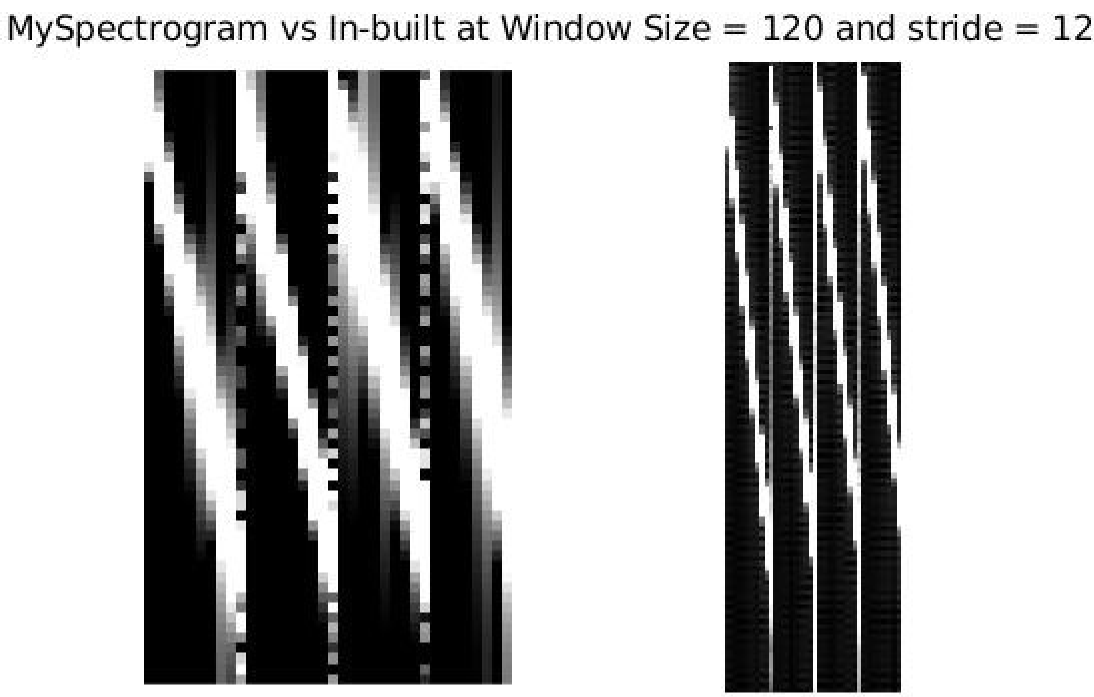
The matrix we get is

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & & \cdots & 0 & 1 \\ & & & & 0 & 1 & 0 \\ \vdots & & & \cdots & 1 & 0 & 0 \\ & & \cdots & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & & & 0 \end{bmatrix}$$

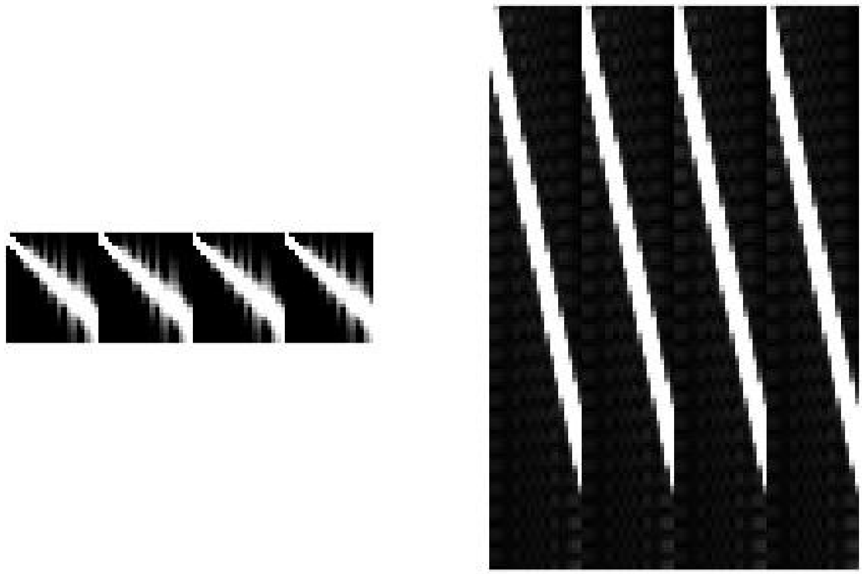
This will flip the image as second row becomes last and last becomes second and so on, therefore, we get inverted image. To restore original image we flip the outp


DOUBLE FFT
FLIP

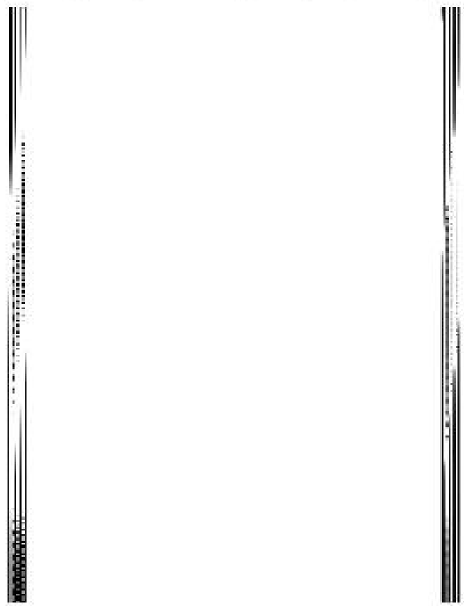
Problem 7 :-


MySpectrogram vs In-built at Window Size = 120 and stride = 12

MySpectrogram vs In-built at Window Size = 50 and stride = 0

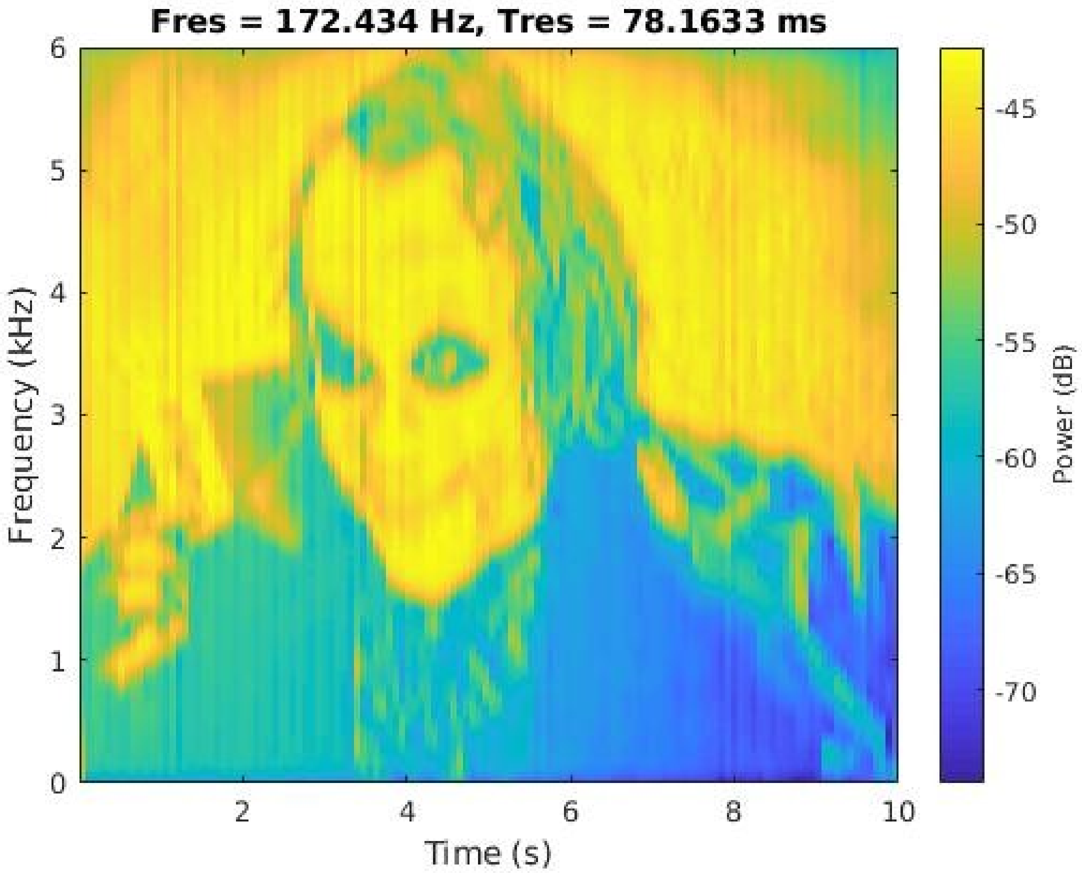


MySpectrogram vs In-built at Window Size = 500 and stride = 50



The size of spectrogram decreases with increasing window size as the no. of discrete samples increase for fixed overlapping. The spread of frequency increases with decreasing the window size. More frequencies are present in a single sample .

(b)
To detect the password , we tried many techniques and found an identifiable image of the joker in the spectrogram and therefore the password is ‘joker’.



(c) The tone has been sampled at frequency 2000 with each tone of size 0.8 sec and gap 0.4sec.