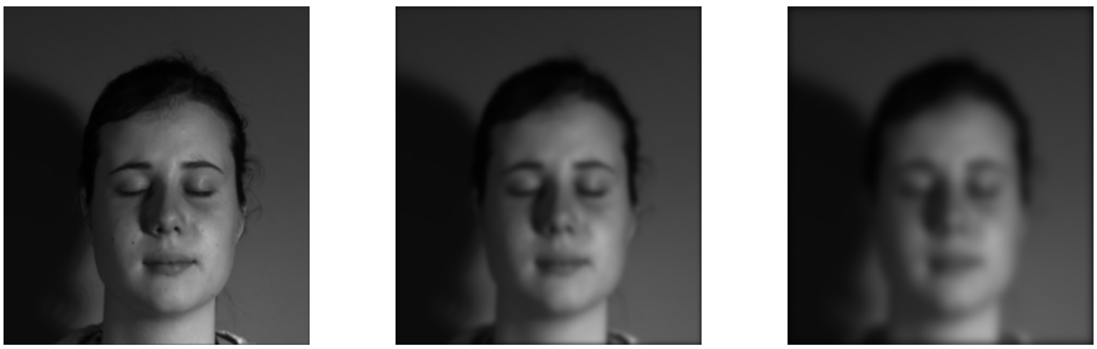
Computational Vision

Task 1: Gaussian Blurring Filter

1. Report the results, describe the effects of larger sigmas, and determine at which step they correspond in biological processing.



(a) (b) (c)

Fig 1: The parameters used here are defined from left to right (a): sigma = 5, size = 30; (b): sigma = 15, size = 90; (c): sigma = 30, size = 180

The application of Gaussian filters with different sigma values to the "face.png" image resulted in distinct effects. Using a sigma of 5 and a filter size of 30 produced a mildly blurred image, maintaining most of the fine details. As the sigma increased to 15 with a filter size of 90, the blurring effect became more pronounced, resulting in a smoother appearance and some loss of fine details. Finally, with a sigma of 30 and a filter size of 180, the image exhibited significant blurring and a softer overall appearance, with a substantial loss of fine details. Larger sigma values in Gaussian filters correspond to biological processing steps where broader-scale information is emphasized while finer details are de-emphasized. This can be related to visual processing mechanisms in the human visual system that integrate visual information across larger receptive fields, enabling the perception of global structures and reducing the impact of noise or small variations in the input.

2. What do you notice about the time taken for each filtering operation and what impact might this have on a complex filter-based model: especially one where filter parameters are optimized by a cost function?

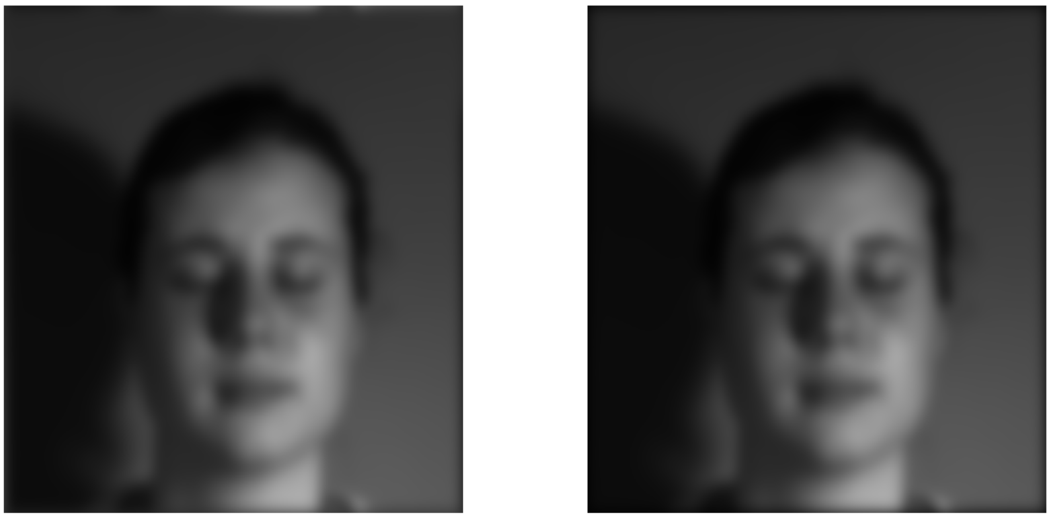
   The time taken for each filtering operation increases with larger filter sizes. In a complex filter-based model with optimized filter parameters, this extended processing time can impact the overall efficiency and speed of convergence. Longer training times may occur, hindering the optimization process and making exploration of the parameter space less efficient. Therefore, the time required for filtering operations plays a crucial role in the performance of complex filter-based models, where filter parameters are optimized through cost function optimization.

3. Looking at the filters using imshow() describe how they compare to retinal receptive fields.

   The Gaussian filters, when visualized using imshow(), show a center-surround organization similar to retinal receptive fields. They have a central region with a high response that gradually decreases towards the surrounding areas. This arrangement resembles the receptive fields in the retina, where cells respond strongly to stimuli in their center and inhibit responses from the surroundings. The Gaussian filters capture local details while blurring the surrounding regions, much like the receptive fields in the retina.

Task 2: Working in Fourier Domain

1. Compare the image produced this way with one produced using conv2(). What differences are there between the filtered images?



(a) (b)

Fig 2: (a) Represents the image produced using the Fourier domain approach and (b) represents the image formed using the conv2 () operation.

When comparing the filtered images produced using the modified Fourier domain approach and the conv2() function, some differences can be observed as seen in Fig 2. The main distinction lies in how the two methods handle the image boundaries during the filtering process. With conv2(), the image is cropped to the same size as the original image, which leads to a slight reduction in size and potential blurring near the boundaries. In contrast, the Fourier domain approach preserves the original image size and accounts for the circular boundaries inherent in the Fourier transform. As a result, the filtered image retains sharper details near the edges. While both methods apply Gaussian filters and exhibit overall blurring effects, the differences in boundary handling can lead to variations in the appearance of the filtered images, particularly near the edges.

2. What are the implications of this for modeling visual processes?

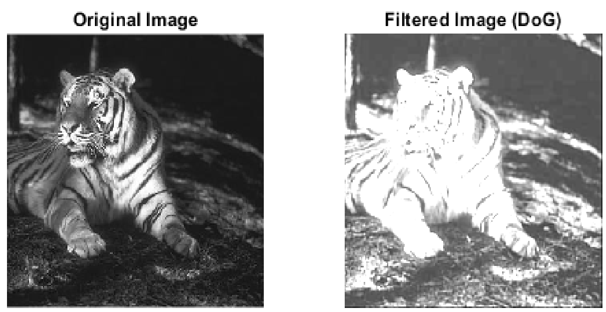
   The differences in boundary handling between the two filtering approaches have implications for modeling visual processes. The cropping operation in conv2() can introduce artifacts near the image boundaries, which may not accurately represent visual processes in the human visual system. On the other hand, the Fourier domain approach preserves the original image size and sharpness near the edges, capturing fine details and edge sensitivity more effectively. This aligns better with the circular nature of visual receptive fields observed in biological systems. However, the Fourier domain approach involves additional computational steps, making it computationally more expensive. Therefore, the choice of filtering approach should consider the trade-off between accuracy and computational efficiency based on the specific requirements of the visual model being developed.

3. What is the possible drawback of the Fourier method?

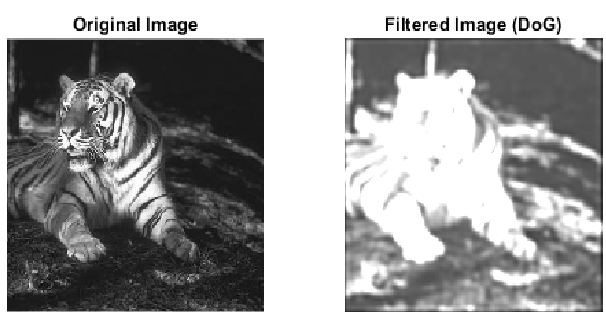
   One possible drawback of the Fourier domain method is the potential introduction of circular artifacts in the filtered image. Due to the periodicity assumption of the Fourier transform, sharp transitions or edges near the image boundaries can result in circular distortions. These artifacts may not accurately represent the original visual information and can have undesired effects in certain applications.

Task 3: Difference of Gaussian (DoG) filters: smoothed edge enhancement

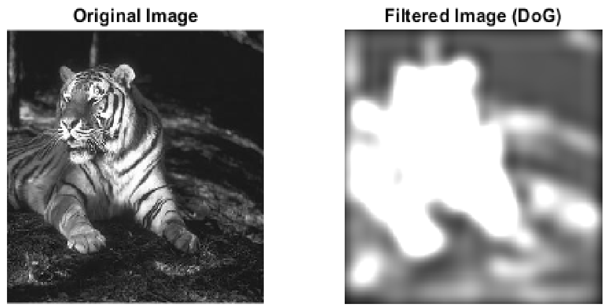
1. What features are picked up by the three filters and what is not picked up? How could you simply improve the situation?



(a)



(b)



(c)

Fig 3: (a) represents the image formed using sigmaPositive as 1 and sigmaNegative as 3, (b) represents the image formed using values 3 and 9, and (c) represents the image formed using values 10 and 30.

The three Differences in Gaussian (DoG) filters with different sigmas will capture different features in the image while leaving out certain details. The filter with a smaller sigma (1, 3) will enhance fine edges and small-scale features but may miss out on larger structures. The filter with medium sigmas (3, 9) will capture a balance between finer and broader edges, but may not emphasize very fine details or extremely large-scale structures. The filter with larger sigmas (10, 30) will focus more on capturing broader edges and larger-scale features, but may not highlight fine details or small-scale edges as prominently. To improve the situation, you can consider using a combination of multiple filters with varying sigmas to capture features across different scales, employing techniques like Gaussian pyramid or adaptive sigma values. These approaches can enhance the ability of the filters to pick up a wider range of features and provide a more comprehensive representation of the image.

2. What is a biologically plausible method to implement filters at different scales?

   A biologically plausible method to implement filters at different scales is by using a multi-scale approach inspired by the receptive field properties of neurons in the visual system. This involves creating a scale-space representation, such as a Gaussian pyramid, which consists of images at different scales obtained through blurring and downsampling. By convolving filters of varying sizes with these images, features can be captured at multiple scales, mimicking the receptive field sizes of neurons in the visual system. This approach aligns with the idea that the visual system processes information at various scales, allowing for the detection of features at different resolutions.

3. In which way a DoG filter mimics retinal processing and what would the sigmas correspond to?

   A DoG filter mimics retinal processing by approximating the receptive field properties of retinal ganglion cells. The DoG filter consists of subtracting the response of a larger-scale Gaussian filter from the response of a smaller-scale Gaussian filter. This creates a center-surround organization, similar to the receptive fields observed in the retina. The sigmas in the DoG filter correspond to the sizes of the Gaussian filters used. The smaller sigma represents the size of the center region, while the larger sigma represents the size of the surrounding region. This center-surround organization and edge enhancement property of the DoG filter closely resemble the receptive field properties of retinal ganglion cells.

4. Referring to images from this workshop, describe the utility of DoG and Gabor filters as models of human vision.

   DoG filters are effective in enhancing edges and capturing edge information, mimicking the receptive field properties of retinal ganglion cells. They are useful for tasks like edge detection and image segmentation. Gabor filters, on the other hand, excel at capturing texture and orientation information. They are inspired by the Gabor wavelets found in the primary visual cortex and are valuable for texture analysis and object recognition.