

به نام خدا

دانشگاه صنعتی امیرکبیر

دانشکده مهندسی کامپیووتر

## پاسخ تمرین سری سوم تصویر پردازی رقمی

استاد:

دکتر رحمتی

دانشجو:

حليمه رحيمى

شماره دانشجویی:

۹۹۱۳۱۰۴۳

بهار ۱۴۰۰

## 1. Fundamentals of 1-D Fourier Transform

a.  $g(t) = \sin(t - 2)$

$$f(t - t_0) \Leftrightarrow F(\omega) e^{-j\omega t_0}$$

$$\mathcal{F}\{\sin(t - 2)\} = \mathcal{F}\{\sin(t)\} e^{-2j\omega}$$

$$\mathcal{F}\{\sin(t - 2)\} = [-j\pi(\delta(\omega - 1) - \delta(\omega + 1))] e^{-2j\omega}$$

b.  $g(t) = \frac{2t^2 + 1}{t^2 + 1}$

$$\mathcal{F}\left\{\frac{2t^2 + 1}{t^2 + 1}\right\} = \mathcal{F}\left\{2 - \frac{1}{1+t^2}\right\} = 4\pi\delta(\omega) - \pi e^{-j\omega t}$$

c.  $g(t) = \delta(t - 1) + \delta(t + 2)$

$$\mathcal{F}\{\delta(t - 1) + \delta(t + 2)\} = e^{-j\omega} + e^{2j\omega}$$

d.  $g(t) = F^{-1}\left\{\frac{1}{f^2 + 1}\right\}$

$$\mathcal{F}\left\{F^{-1}\left\{\frac{1}{f^2 + 1}\right\}\right\} = \frac{1}{f^2 + 1}$$

e.  $g(t) = \omega t e^{-|t|}$

$$\mathcal{F}\{\omega t e^{-|t|}\} = \frac{2\omega^2 + 2}{\omega^4 + 4}$$

$$f. g(t) = \cos(4\pi t) \cos(6\pi t)$$

$$\Im \{ \cos(4\pi t) \cos(6\pi t) \}$$

$$= \frac{1}{2}\pi [\delta(\omega - 10\pi) + \delta(\omega - 2\pi) + \delta(\omega + 2\pi) + \delta(\omega + 10\pi)]$$

$$g. \operatorname{Re} \{ X(j\omega) \} = 0$$

لأن  $g$  معرفة في  $\omega$  Real  $\Rightarrow$   $x(t)$

$$h. \operatorname{Im} \{ X(j\omega) \} = 0$$

لأن  $X(j\omega)$  معرفة في  $\omega$  Real  $\Rightarrow$   $x(t)$

$$i. \int_{-\infty}^{\infty} X(j\omega) d\omega = 0$$

لأن  $\int_{-\infty}^{\infty} X(j\omega) d\omega = \int_{-\infty}^{\infty} \operatorname{Re} \{ X(j\omega) \} d\omega = 0$

لأن  $i$  هي  $\operatorname{Im} \{ X(j\omega) \} = 0$

$$j. \int_{-\infty}^{\infty} \omega X(j\omega) d\omega = 0$$

لأن  $\int_{-\infty}^{\infty} \omega X(j\omega) d\omega = \int_{-\infty}^{\infty} \operatorname{Re} \{ X(j\omega) \} \omega d\omega = 0$

K.  $X(j\omega)$  is periodic.

برای کسری  $k, h$  نویسید

1. There is a real  $\alpha$  such that  $e^{j\alpha\omega} X(j\omega)$  is real.

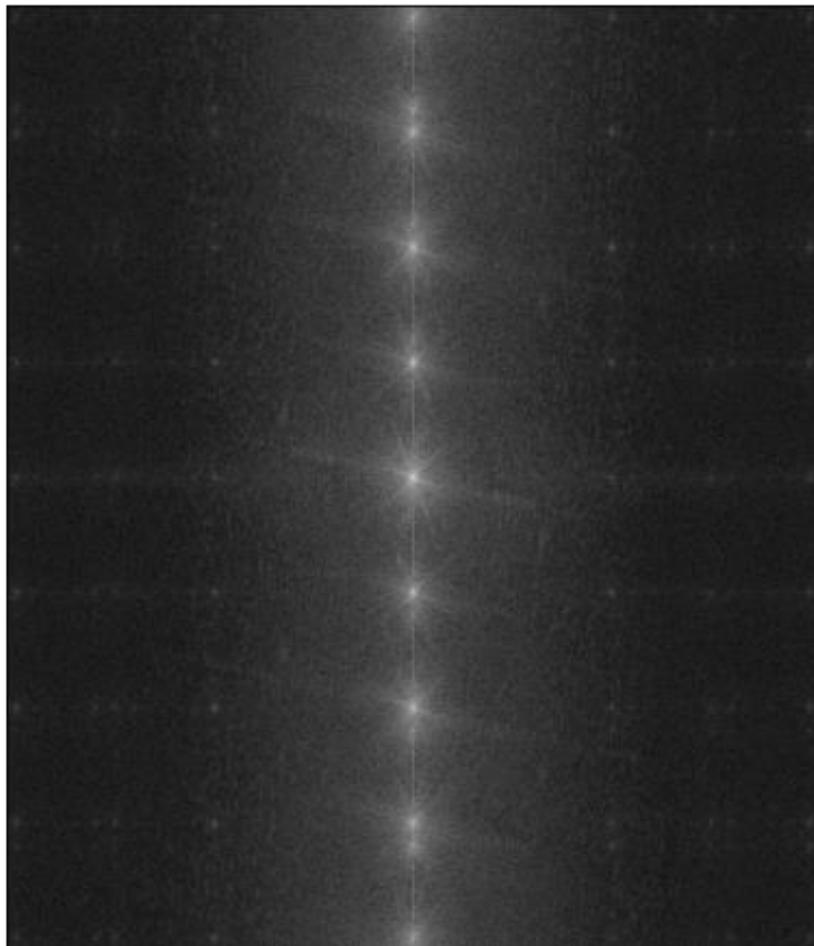
لذا  $j\omega + \alpha$  باید مولفه ای حقیقی باشد

## 2. The Mysterious Moiré Pattern

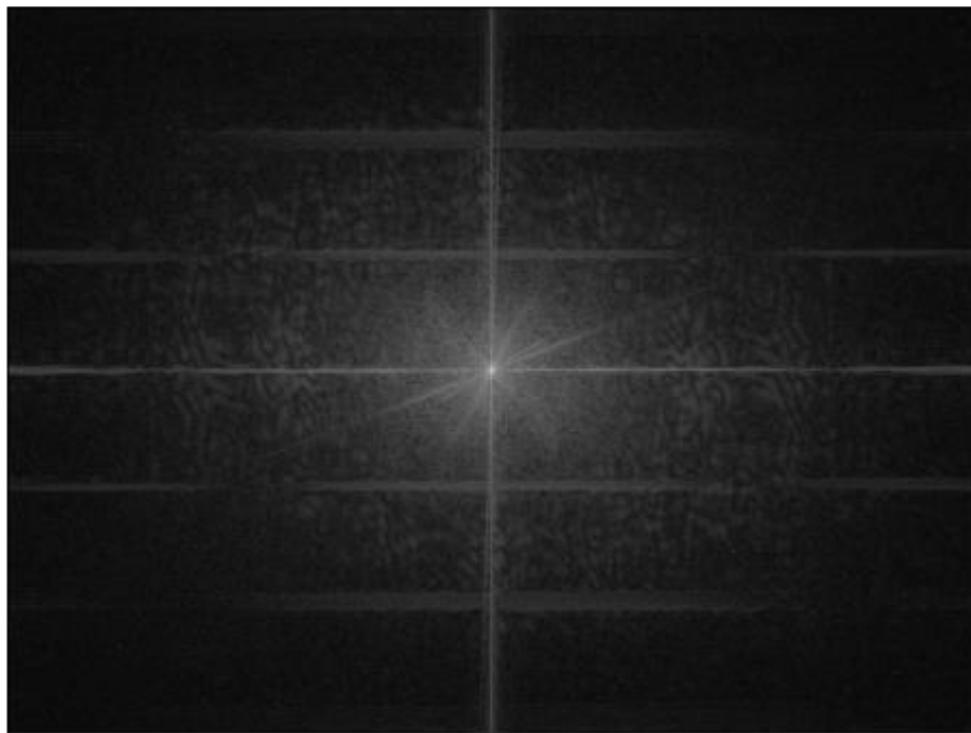
I

- a. Display the images spectrum in the frequency domain. Give your interpretations of the resultant spectrum images.

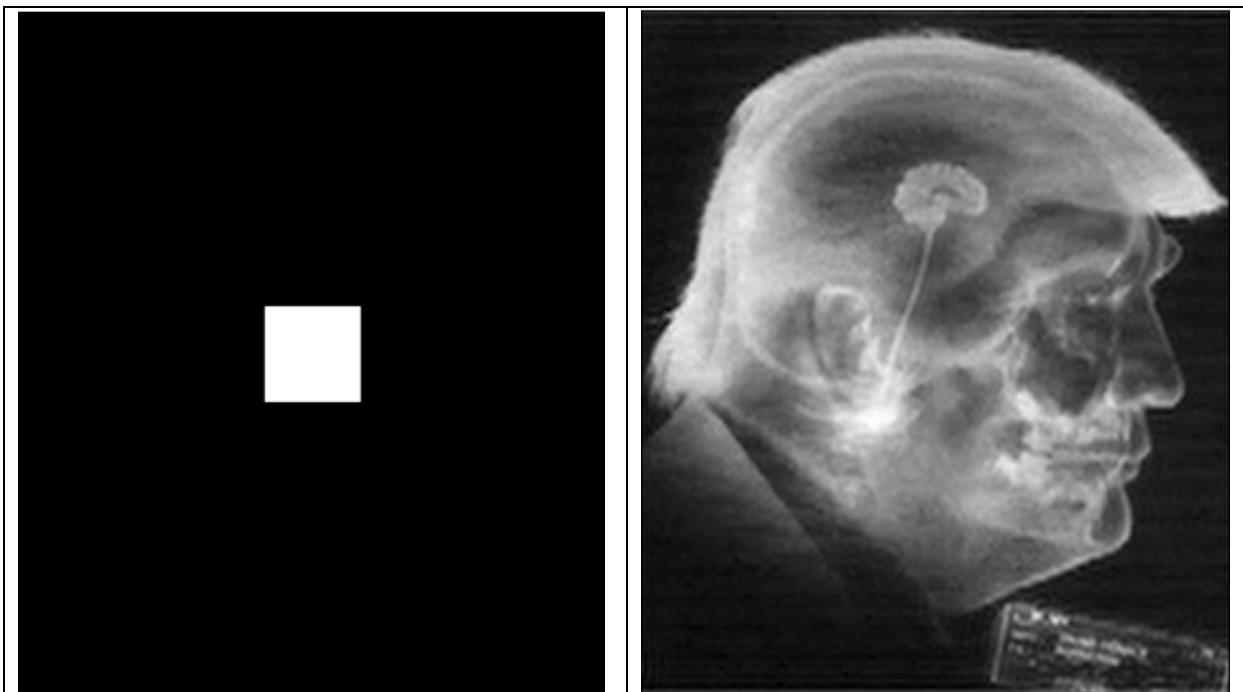
برای تصویر x-ray

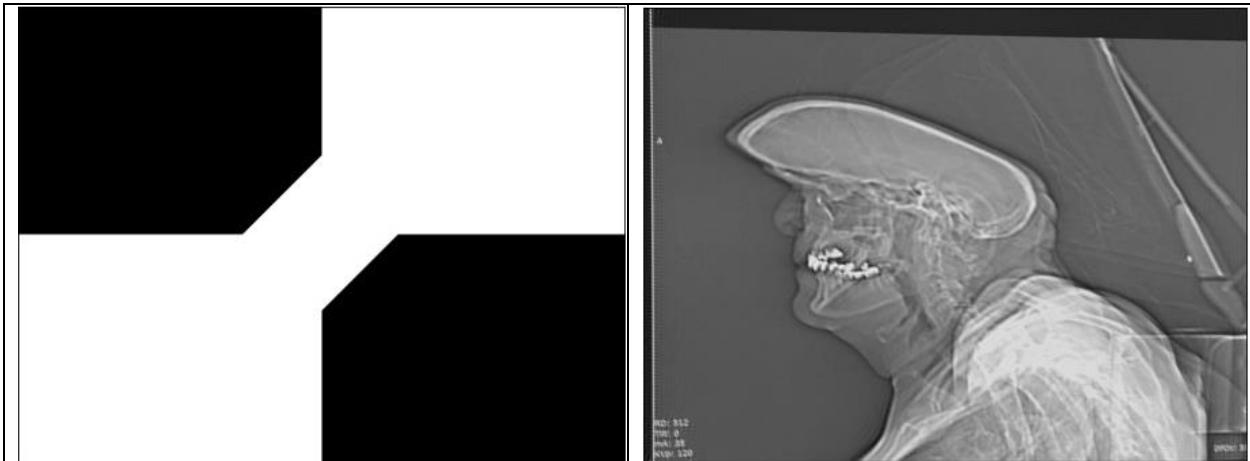


برای تصویر CT scan



- b. Design appropriate filters and apply them on the images to reduce the effect of Moiré pattern. Display the results as well as the filters used.

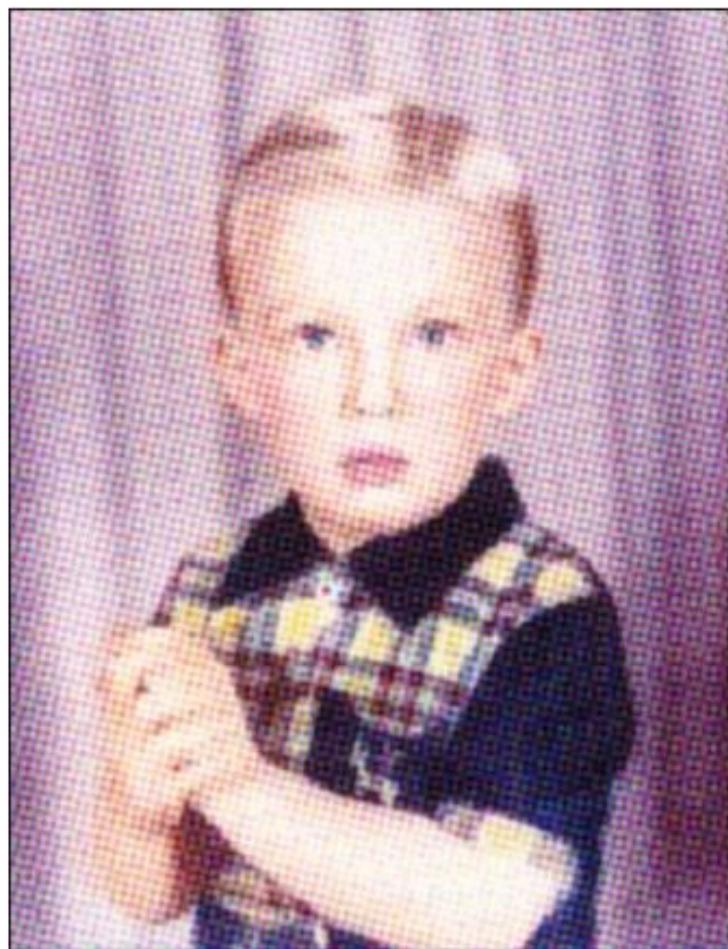


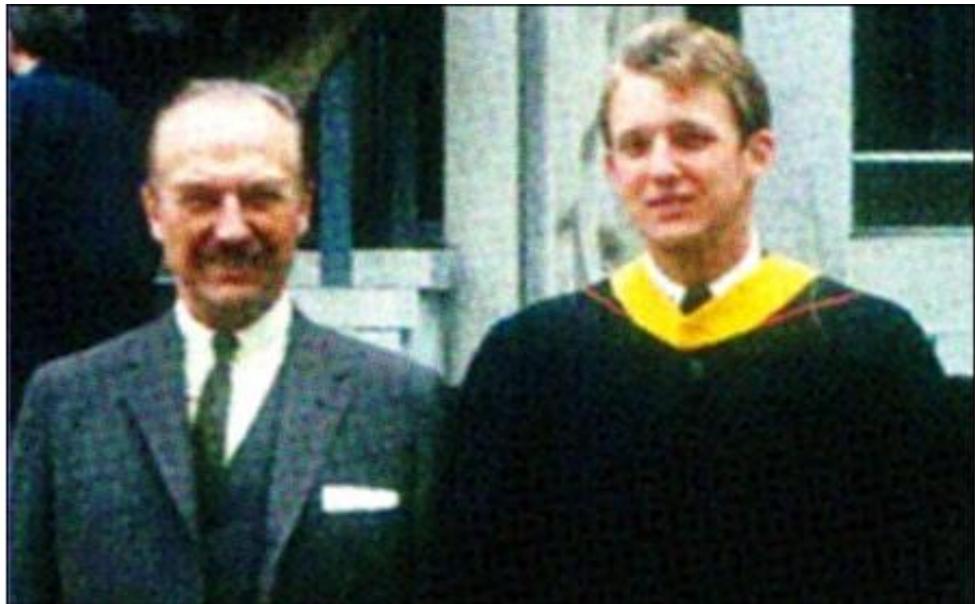


می شود برای تصویر اولی فقط بخش وسط را به صورت افقی نگاه داشت، تقریباً نتیجه یکسان می دهد.  
خطوط نویز در جهتی هستند که در ماسک سیاه کرده ام.

II

c. Implement a method in spatial domain to reduce the halftone effect.

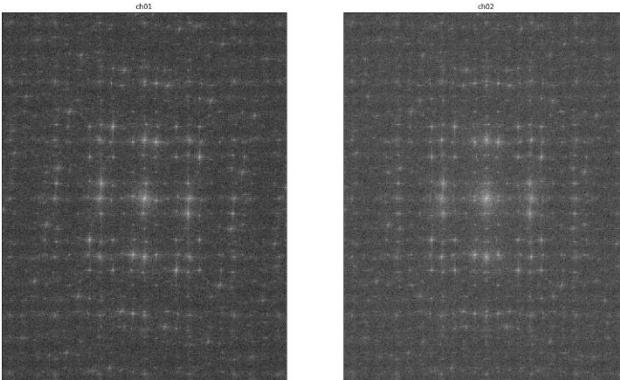
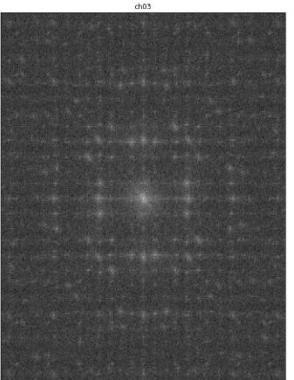


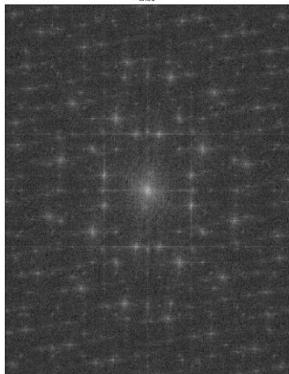
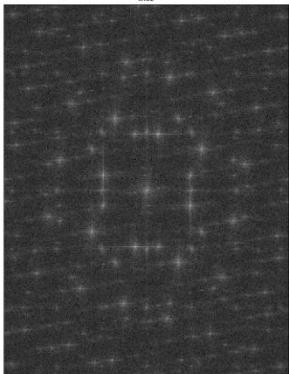
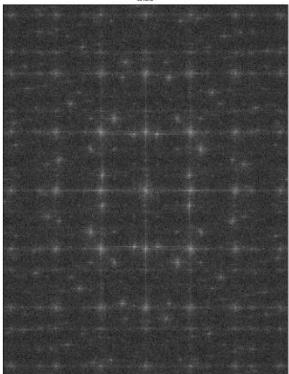
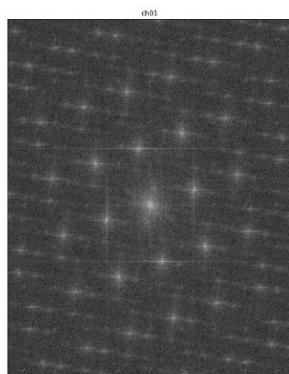
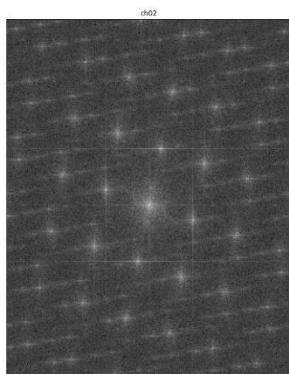
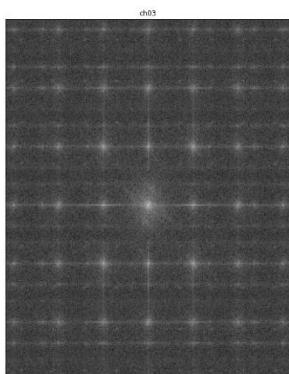
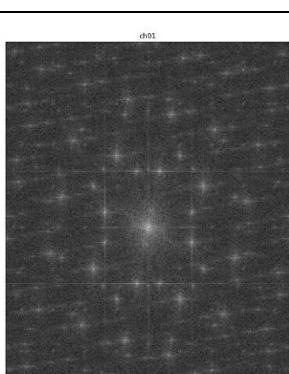
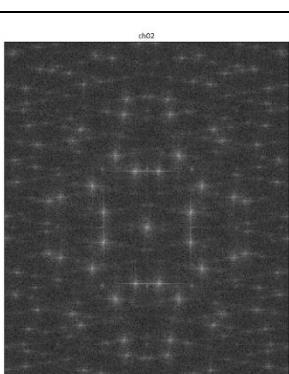
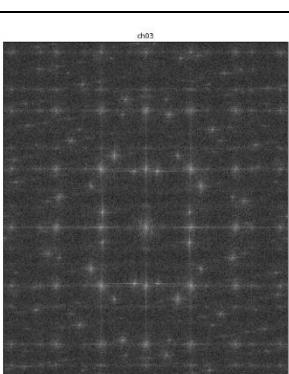


فیلتر  $5 \times 5$  میانگین گیرنده را استفاده کردم.

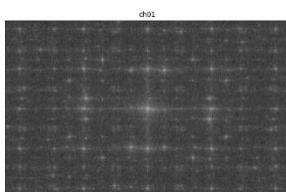
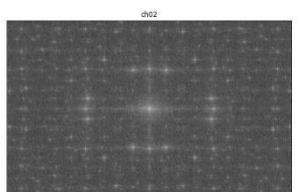
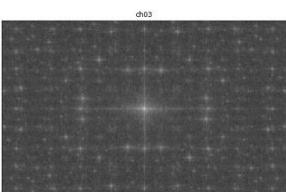
d. Now convert the given RGB images to different color spaces, namely **HSV**, **YCbCr** and **Lab**. Display the Fourier transform of each channel (R, G, B, H, S, V, Y, Cb, Cr, L, a and b) and determine which color space the Fourier transform is more informative in.

به ترتیب چهار تای اول برای تصویر کودکی و چهار تای بعدی برای تصویر فارغ التحصیلی است. همچنین به ترتیب HSV، Lab، YCbCr و YCrCb می باشند که نتیجه‌ی کانال‌های آنها به ترتیب از چپ به راست قرار گرفته‌اند.  
تصویر کودکی

HSV		
-----	--	---

Lab			
RGB			
YCrCb			

تصوير فارغ التحصيلي

HSV			
-----	---	--	---

Lab			
RGB			
YCrCb			

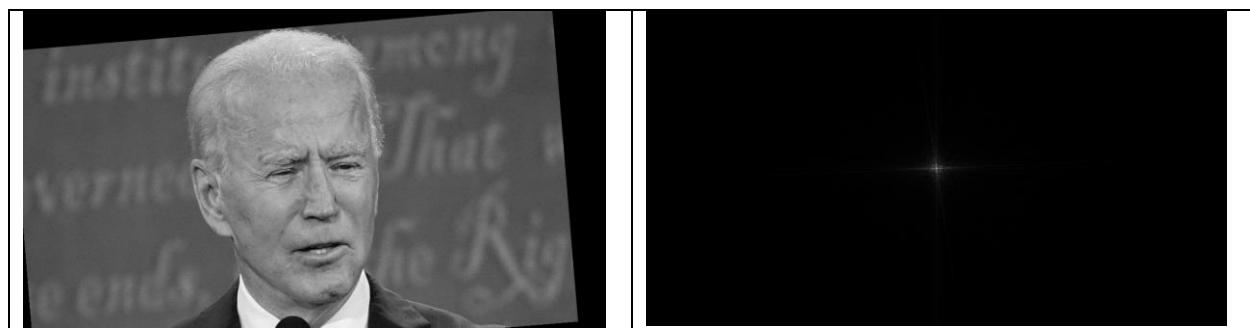
به نظر می رسد RGB بپردازی این نتایج را که نشاندهند نویز هستند تشخیص داد و ریز و پراکنده نیستند. می توان مشاهده کرد که گویی مقادیر چرخش زاویه ای نسبت به هر کanal RGB داشته اند و کanal سبز و آبی در زوایای مخالف هم اند.

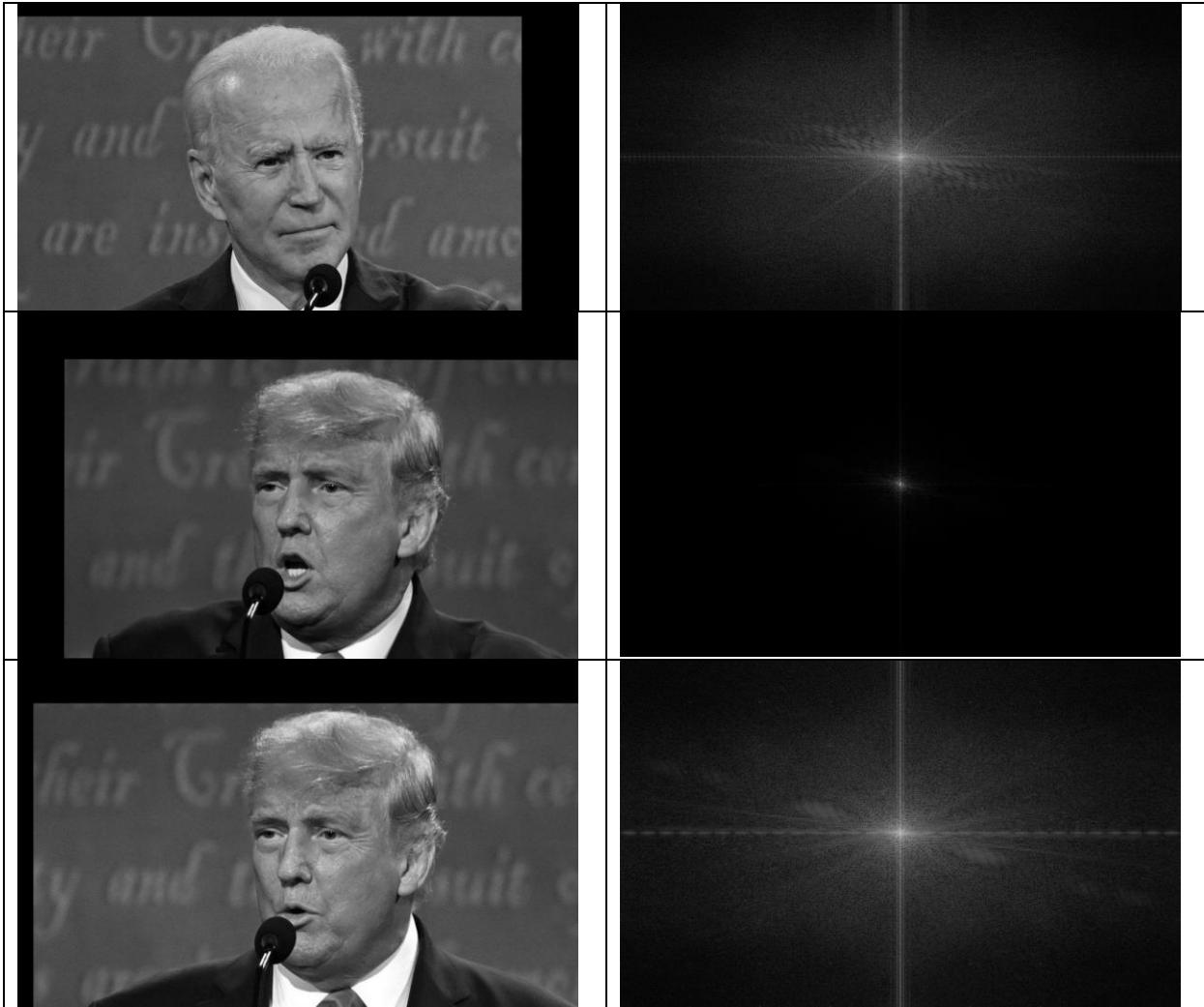
e. Use the color space determined in the previous part, and design suitable filters for each channel separately to reduce the halftone effect. Combine the filtered channels to obtain filtered color results. Display the resultant images as well as the filters.

f. Compare the results obtained in part (e) with part (c).

### 3. Implementing the Bittersweet emotion in Photos!

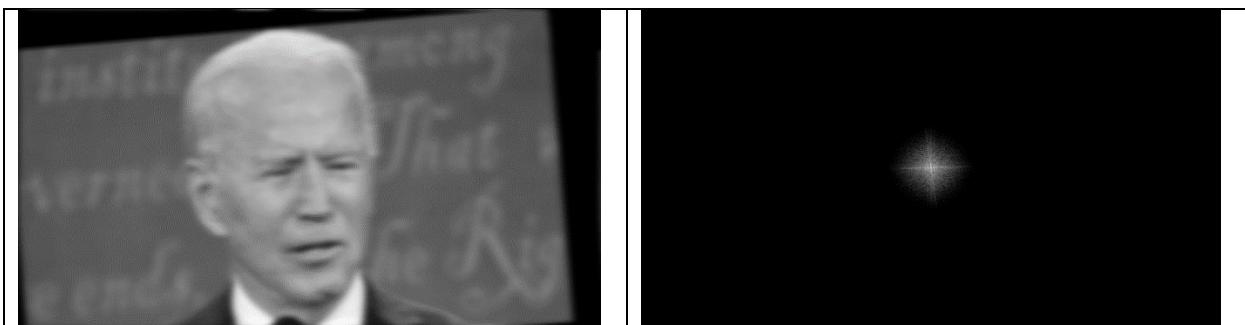
a. Image alignment. A function (`align_imgs.m` for MATLAB and `align_imgs.py` for Python) is provided for you, which takes two images and two pairs of points, and align them so that the two pairs of points will have approximately equal coordinates. Use this function to align the input images. Display the resultant aligned images and their amplitudes of the Fourier transform.

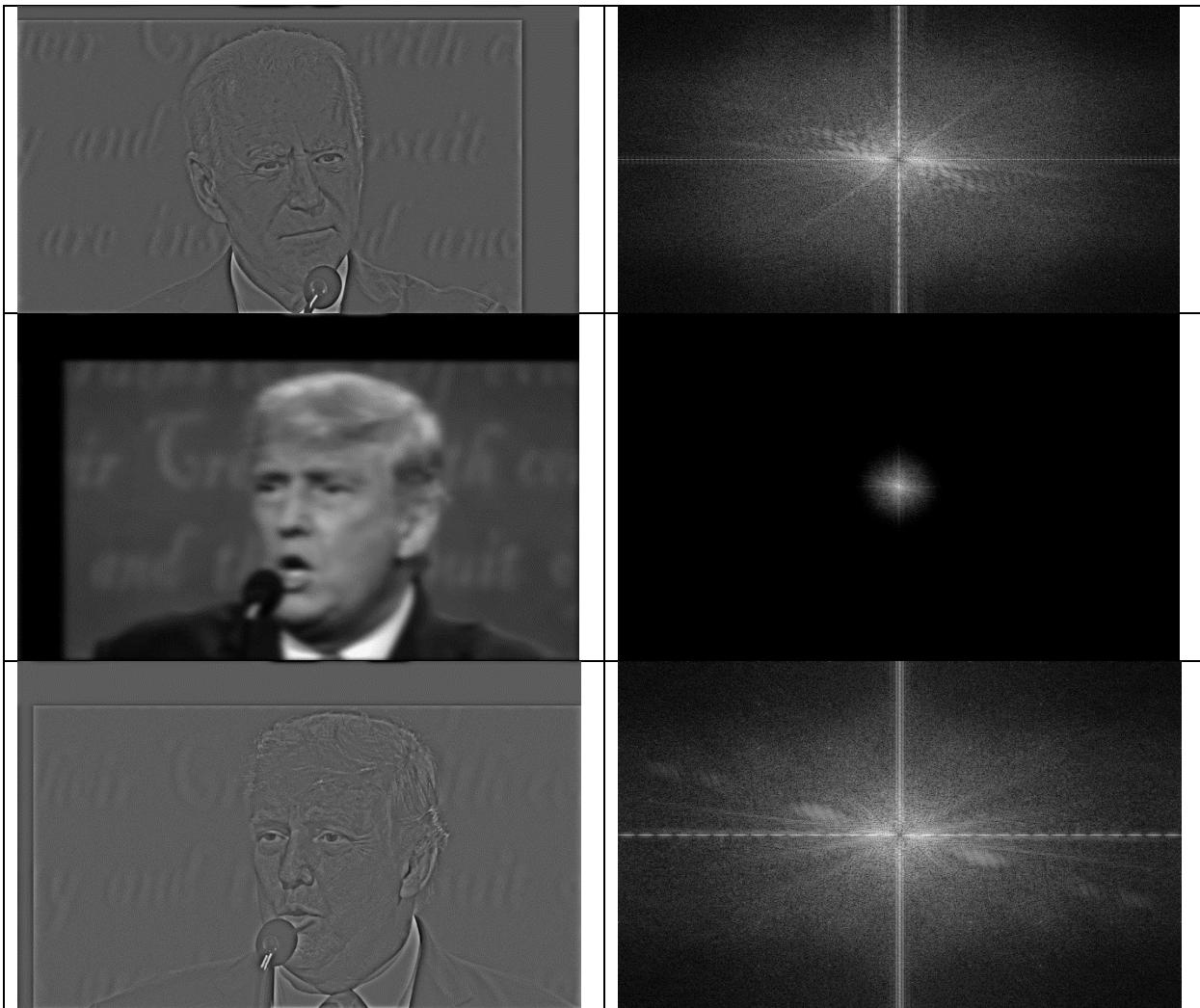




دو نقطه را وسط چشم‌های افراد در نظر گرفتم؛ چراکه اصولاً فاصله‌ی چشم‌ها تغییر نمی‌کند.

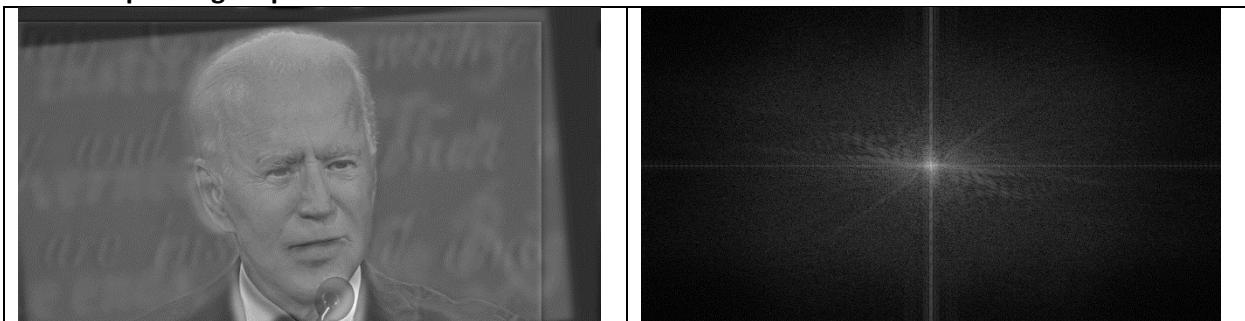
**b. Image filtering.** Apply low-pass filter on the first image using a standard Gaussian filter, and high-pass filter on the other by subtracting the image filtered with Gaussian filter from the original one. Choose proper values for cut-off frequencies. Display the results as well as their logarithmic amplitude of the Fourier transform.

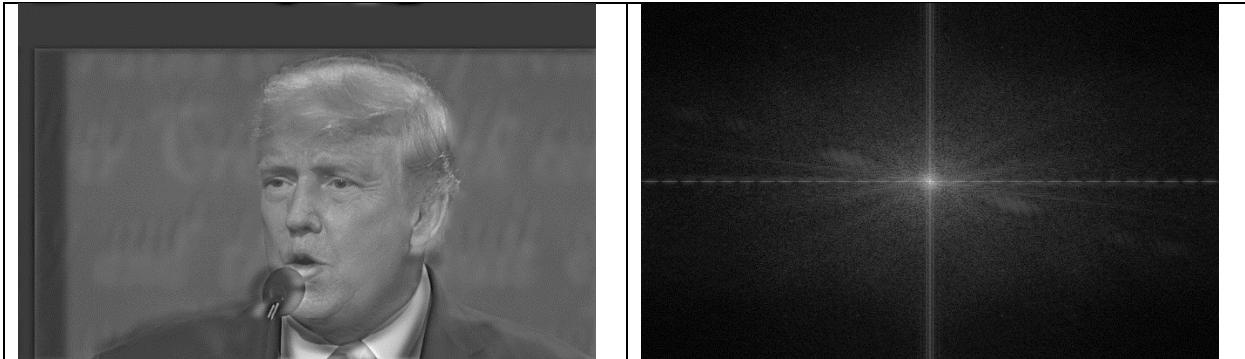




مقدار cutoff-frequency در اینجا برابر با ۲۵ می باشد.

c. Merge Images. Merge the images you obtained in the previous part, and display the final image and the corresponding amplitude of the Fourier transform.





**d. Visualization.** Apply a Gaussian filter with five increasing cut-off values on the resultant images in order to illustrate the process of transformation of one expression into another.

مقادیر cutoff-frequency به ترتیب از چپ به راست برابر با ۵، ۱۵، ۲۵، ۳۵ و ۴۵ می‌باشد.

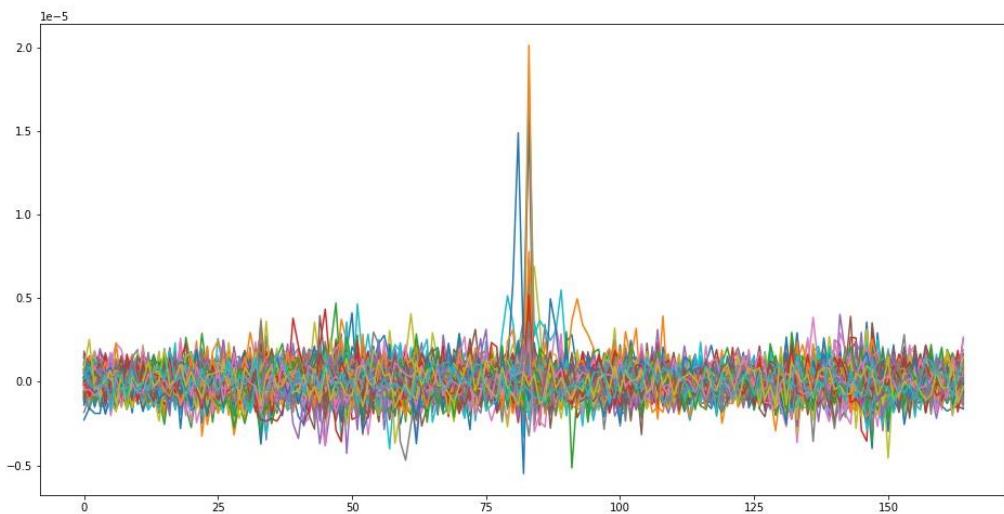




#### 4. Who is Who: Face Recognition in the Frequency Domain

a. Ignore the images with occlusions, and keep the images with postfix 01 to 07 as the training images, and those with postfix 14 to 20 as the test images. Pick a random test image, and perform face recognition using the above algorithm. Display the MACE filter, output correlation and the recognition result.





عکس تست دوم از شخص دوم ('M-002-15') را برای تست دادم و نتیجه را به شکل دو بعدی نمایش دادم. هر یک از ستون‌های نتیجه‌ی ضرب تصویر در فیلتر، یک خط در پلات بالا می‌باشد (تصویر ۱۲۰\*۱۶۵ می‌باشد). پیک بهوضوح قابل مشاهده است. الگوریتم شخص را به درستی تشخیص داده است. برای تعیین خودکار نتیجه، از فرمولی که در بخش b نوشتم استفاده کرده‌ام.

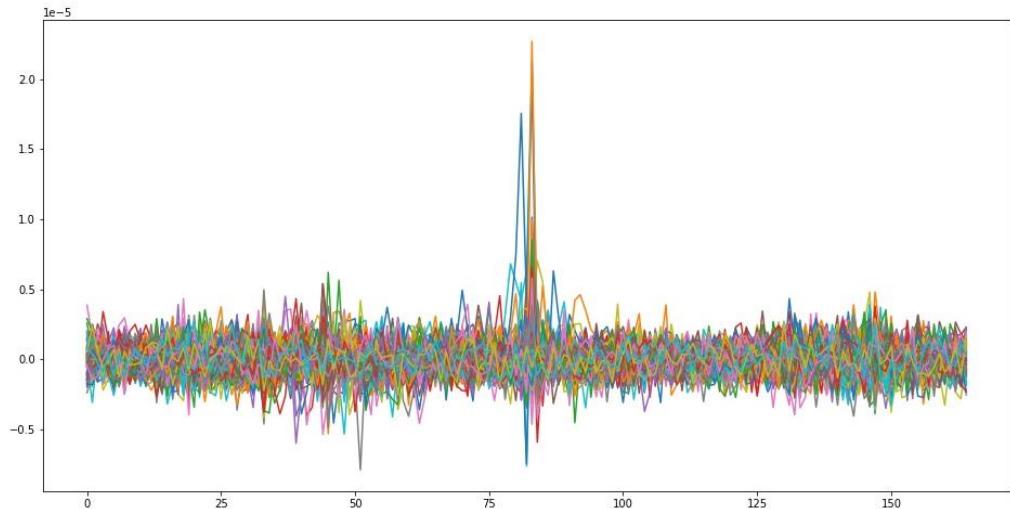
**b. Calculate the recognition rate by dividing the number of true identifications to total number of test images.**

درجه شناسایی برابر با ۷۱ درصد شد. در تصویر درون فolder *\_b* نتیجه‌ی شناسایی هر تست و درجه شناسایی را برای هر کلاس مشاهده می‌کنید. به علاوه به علت حجم زیاد، تنها برخی از نتایج کلاس‌ها را در فolder قرار داده ام تا میزان تشخیص فرمول به کار رفته (که عملکرد معقولی داشته) را نمایش داده باشم.  
برای آنکه به طور خودکار برنامه پیک را تشخیص دهد و بسنجد، از آنچه در تحقیقات پیشین<sup>۱</sup> استفاده شده بود، یعنی Peak to sidelobe ration (PSR) بهره برمد. به این صورت که یک بخش ۲۰ در در وسط خروجی ضرب را در نظر گرفتم و بیشترین مقدار را در آن به عنوان پیک انتخاب کدم، سپس با  $\frac{p-\mu}{\sigma}$  PSR = و ترشولد ۸، به سنجش تشخیص برنامه پرداختم.

**c. Now consider the whole dataset, i.e. images with postfix 01 to 13 as the training images, and those with postfix 14 to 26 as the test images. Pick a random test image, and perform face recognition using the above algorithm. Display the MACE filter, output correlation and the recognition result.**



<sup>1</sup> Face Verification using Correlation Filters,  
<https://cedar.buffalo.edu/~govind/CSE717/papers/Face%20Verification%20using%20Correlation%20Filters.pdf>



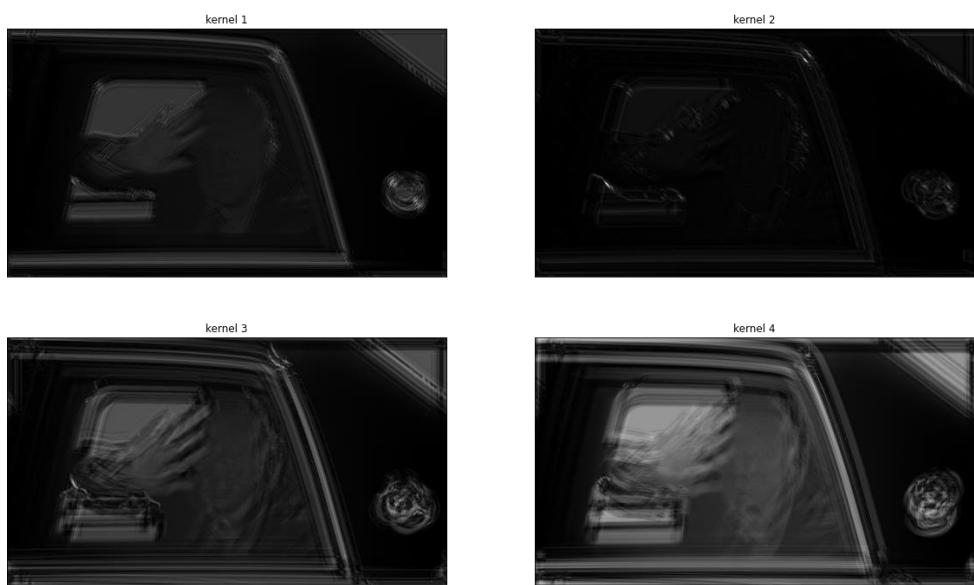
همچنان از همان تصویر استفاده کردم تا تفاوت‌ها قابل مشاهده باشد. نتیجه تقریباً شبیه است و تفاوت‌های بسیار کم آن به دلیل اضافه کردن تصاویر جدید به مجموعه آموزش می‌باشد.  
لازم به ذکر است تمامی فیلترها در فولدرهای مربوط به خودشان قرار گرفته‌اند.

#### d. Calculate the recognition rate, and compare the result with part (b).

درجه شناسایی برابر با ۶۸ درصد شد، در تصویر درون فولدر `part_d` نتیجه‌ی شناسایی هر تست و درجه شناسایی را برای هر کلاس مشاهده می‌کنید. به علاوه به علت حجم زیاد، تنها برخی از نتایج کلاس‌ها را در فولدر قرار داده ام تا میزان تشخیص فرمول به کار رفته (که عملکرد معقولی داشته) را نمایش داده باشم.  
با توجه به آنکه اکنون تصاویر آموزش و آزمون دارای تصاویر `occluded` می‌باشد، درجه‌ی شناسایی کمتر از بخش b قابل درک است.

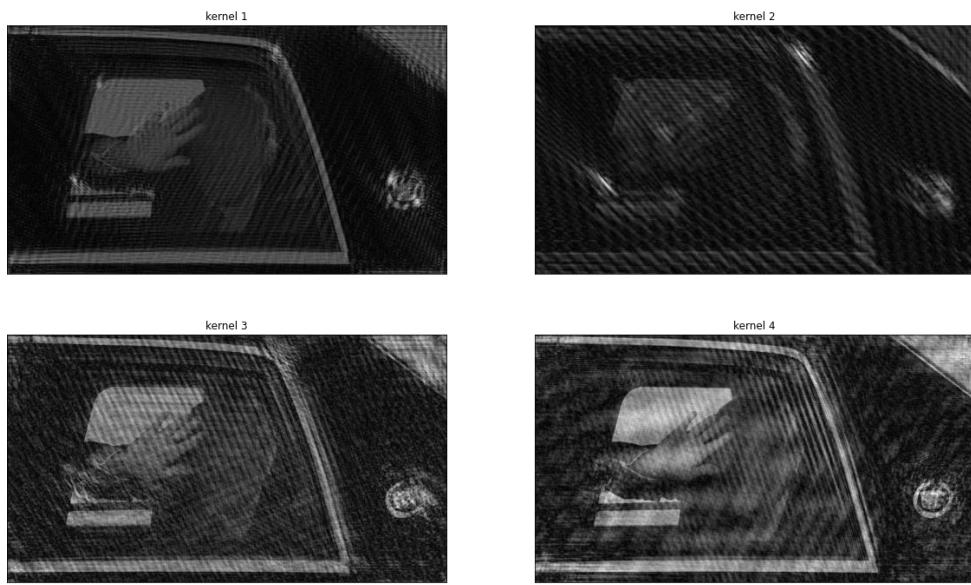
### 5. Stop Moving: Dealing with the Motion Blur

#### a. Use the PSF kernels and blur the input images by performing image convolution in the frequency domain. Display the results (8 in total).



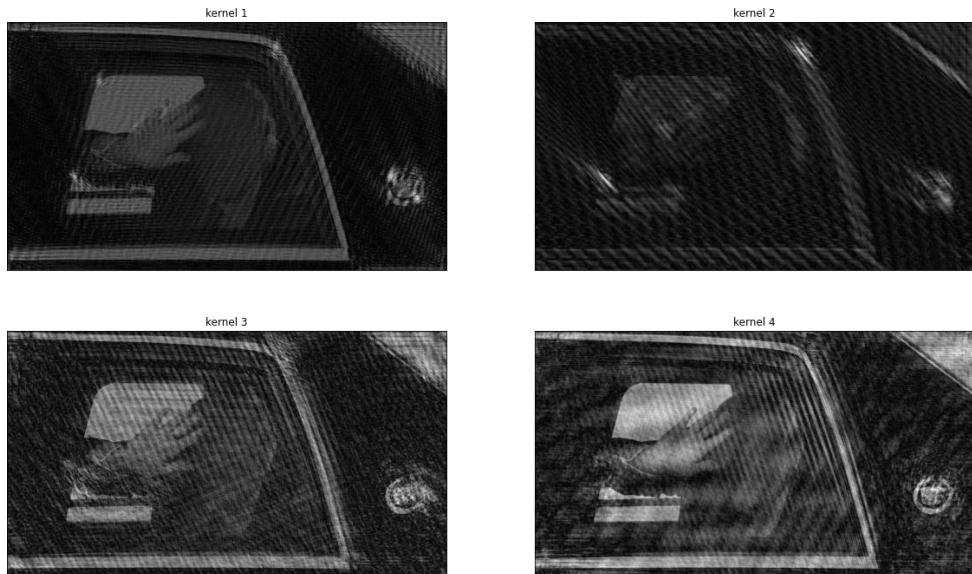


**b. Use the least-square problem (Equation (4)) and recover the input images using blurred images and their corresponding PSF kernels.**





c. Use the least-square problem (Equation (6)) with  $A = I$ , and recover the input images using blurred images and their corresponding PSF kernels. Evaluate the deconvolution result using quality measurements such as PSNR and SSIM.



```
PSNR donald_1_deblur_tls_kernel1 6.79860458134789
PSNR donald_1_deblur_tls_kernel2 6.810627847908122
PSNR donald_1_deblur_tls_kernel3 6.797744479910566
PSNR donald_1_deblur_tls_kernel4 6.7950844821658265
```



```

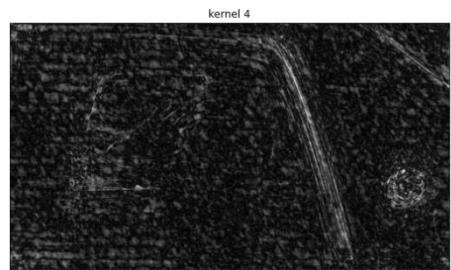
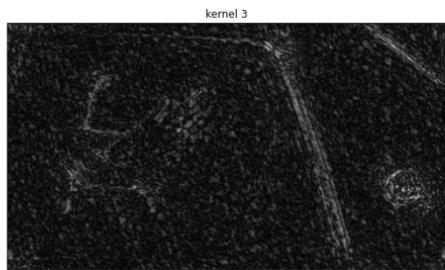
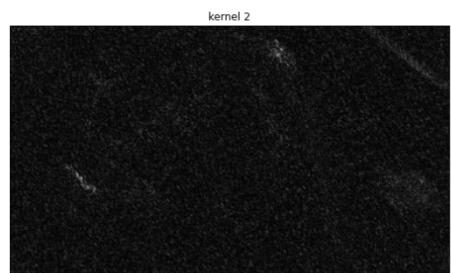
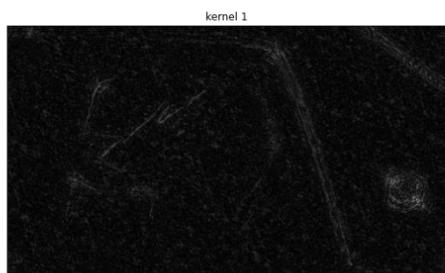
PSNR donald_2_deblur_tls_kernel 1_channel 1 11.790384902670986
PSNR donald_2_deblur_tls_kernel 1_channel 2 12.216228649371452
PSNR donald_2_deblur_tls_kernel 1_channel 3 12.179509212394146
PSNR donald_2_deblur_tls_kernel 2_channel 1 9.789672921477678
PSNR donald_2_deblur_tls_kernel 2_channel 2 9.890246511648051
PSNR donald_2_deblur_tls_kernel 2_channel 3 9.773857793878047
PSNR donald_2_deblur_tls_kernel 3_channel 1 13.000947188663952
PSNR donald_2_deblur_tls_kernel 3_channel 2 13.222749643763205
PSNR donald_2_deblur_tls_kernel 3_channel 3 12.86797463099967
PSNR donald_2_deblur_tls_kernel 4_channel 1 16.080360810520475
PSNR donald_2_deblur_tls_kernel 4_channel 2 16.166711158841874
PSNR donald_2_deblur_tls_kernel 4_channel 3 15.996960963145712

```

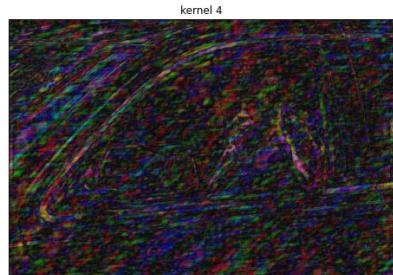
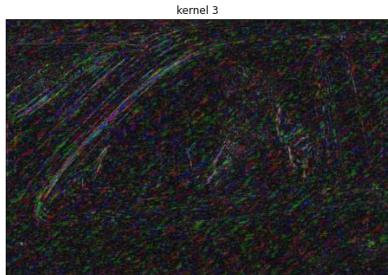
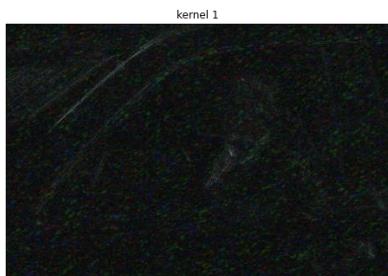
تصاویر مشابه قبل اند، با بزرگتر شدن لاندا، مقادیر بیشتر کنترل می شوند و موجب می شود تغییرات در اینجا به خوبی قبل نباشد. البته بهتر از حالی بود که هیچ عملی بر روی آن انجام نشده. در واقع پاسخ بخش پیشین برای تصویر ما مناسب بوده و نیازی به اضافه کردن ماتریس A و ضریب لاندا نبود. لاندا را برابر با 0.002 در نظر گرفتم.

**d. Use the least-square problem (Equation (6)) with  $A = D$ , where  $D$  is a Laplacian operator, and recover the input images using blurred images and their corresponding PSF kernels. Evaluate the deconvolution result using quality measurements such as PSNR and SSIM.**

مقدار PSNR کم شده و نتیجه به خوبی قبل نیست. لاندا را تغییر ندادم.



PSNR donald\_1\_deblur\_tls\_kernel1 6.774585058558084  
PSNR donald\_1\_deblur\_tls\_kernel2 6.774463751552256  
PSNR donald\_1\_deblur\_tls\_kernel3 6.774898321109127  
PSNR donald\_1\_deblur\_tls\_kernel4 6.775101255731042



```

PSNR donald_2_deblur_tls_kernel 1_channel 1 9.53286881597298
PSNR donald_2_deblur_tls_kernel 1_channel 2 9.726812455011293
PSNR donald_2_deblur_tls_kernel 1_channel 3 9.377208346781199
PSNR donald_2_deblur_tls_kernel 2_channel 1 9.893475810160515
PSNR donald_2_deblur_tls_kernel 2_channel 2 9.94083946859568
PSNR donald_2_deblur_tls_kernel 2_channel 3 9.791860822519519
PSNR donald_2_deblur_tls_kernel 3_channel 1 10.145883491244676
PSNR donald_2_deblur_tls_kernel 3_channel 2 10.200801103450925
PSNR donald_2_deblur_tls_kernel 3_channel 3 10.029185490652939
PSNR donald_2_deblur_tls_kernel 4_channel 1 10.571183432546283
PSNR donald_2_deblur_tls_kernel 4_channel 2 10.465538623296275
PSNR donald_2_deblur_tls_kernel 4_channel 3 10.300238682347187

```

e. Rewrite a new optimization problem similar to (5) to obtain unknown PSF kernel given a clean image as well as its corresponding blurry version. Write down its closed-form solution, and use it to obtain PSF kernels when the original input images are given. Evaluate the results using quality measurements such as RMSE.

از آن جهت که  $H$  و  $x$  به صورت عنصر به عنصر در یکدیگر ضرب شده اند، کافیست در نتیجه نهایی به جای  $H$ ،  $x$  قرار دهیم.  
با این حال نحوه محاسبه را در زیر می بینید:

$$\hat{H} = \underset{H}{\operatorname{argmin}} \|y - Hx\|_2^2 + \lambda \|Ax\|_2^2$$

This time we do have  $y$  and  $x$ .

$$(y - Hx)^T (y - Hx) + \lambda (Ax)^T (Ax)$$

$$= (y^T - x^T H^T) (y - Hx) + \lambda (H^T A^T A H)$$

$$= y^T y - x^T H^T y - y^T H x + x^T H^T H x + \lambda x^T A^T A x$$

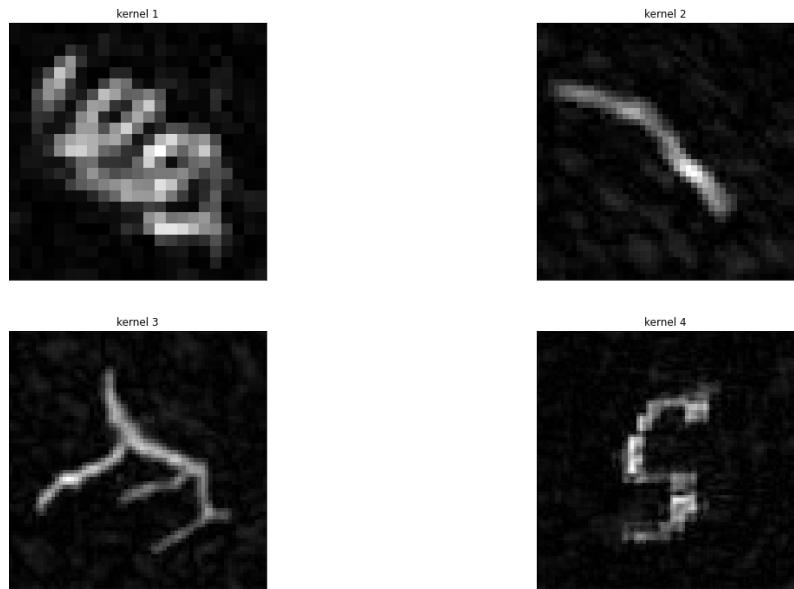
$$= y^T y - 2 x^T H^T y + x^T H^T H x + \lambda x^T A^T A x$$

$$\Rightarrow -2 x^T H^T y + x^T H^T H x + \lambda x^T A^T A x$$

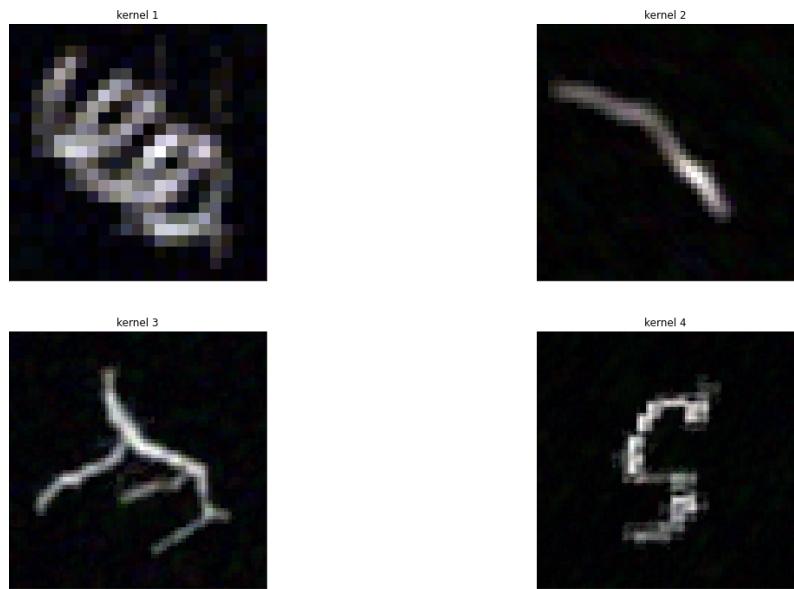
$$\Rightarrow x^T H + \lambda A^T A x = 2 x^T y$$

$$\Rightarrow H = 2(x^T + \lambda A^T A)^{-1} x^T y$$

$$\text{we can ignore (2)} \Rightarrow \hat{H} = (x^T + \lambda A^T A)^{-1} x^T y$$



```
RMSE donald_1.blur.kernel1 66.4048900262595
RMSE donald_1.blur.kernel2 26.242386064550203
RMSE donald_1.blur.kernel3 42.16646896448953
RMSE donald_1.blur.kernel4 37.24775903655373
```



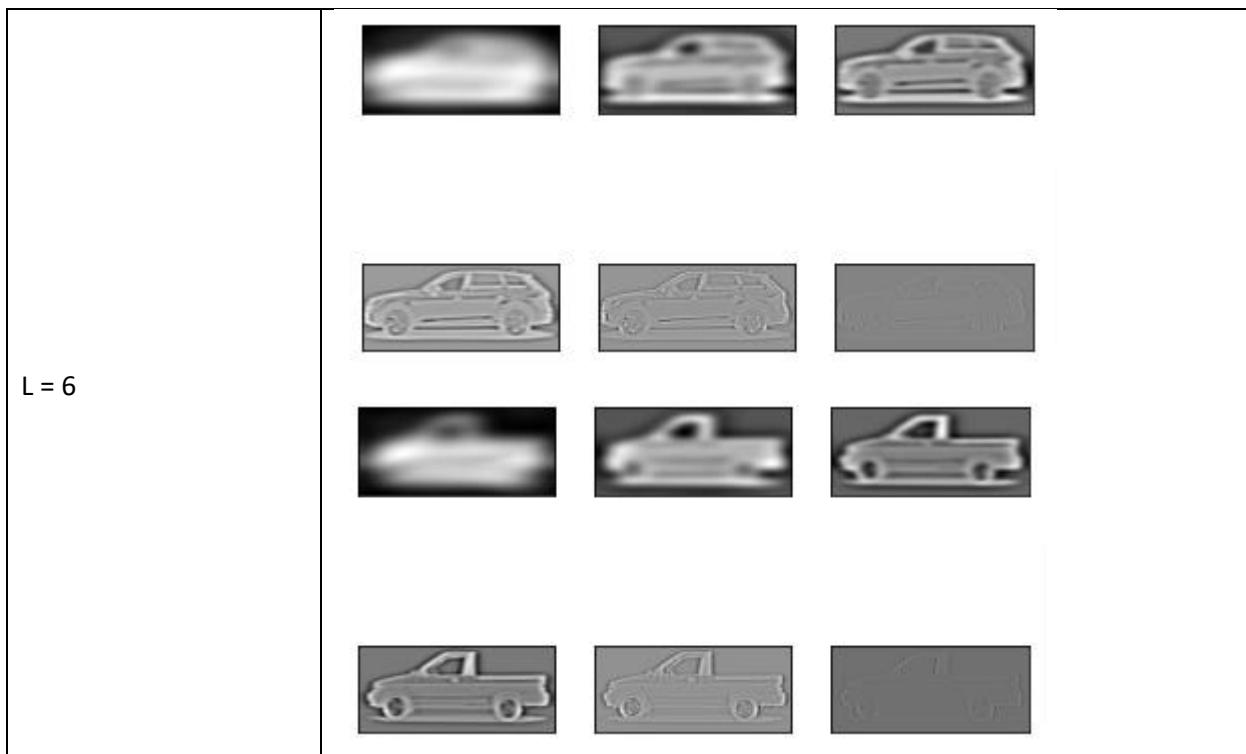
```
RMSE donald_2.blur_kernel 1_channel 1 12.849514108681015
RMSE donald_2.blur_kernel 1_channel 2 5.3728529452013
RMSE donald_2.blur_kernel 1_channel 3 5.319974398155413
RMSE donald_2.blur_kernel 2_channel 1 6.870211874078219
RMSE donald_2.blur_kernel 2_channel 2 6.944680374486666
RMSE donald_2.blur_kernel 2_channel 3 7.685846332810047
RMSE donald_2.blur_kernel 3_channel 1 9.302931465025482
RMSE donald_2.blur_kernel 3_channel 2 6.499503616350636
RMSE donald_2.blur_kernel 3_channel 3 6.124729079026753
RMSE donald_2.blur_kernel 4_channel 1 5.374641158267251
RMSE donald_2.blur_kernel 4_channel 2 4.931340181398711
RMSE donald_2.blur_kernel 4_channel 3 3.7428208197638155
```

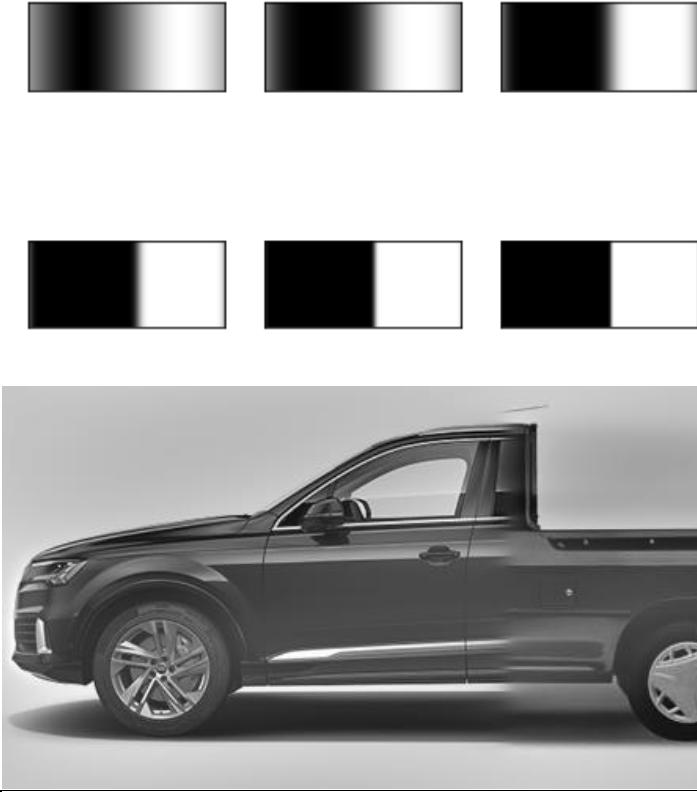
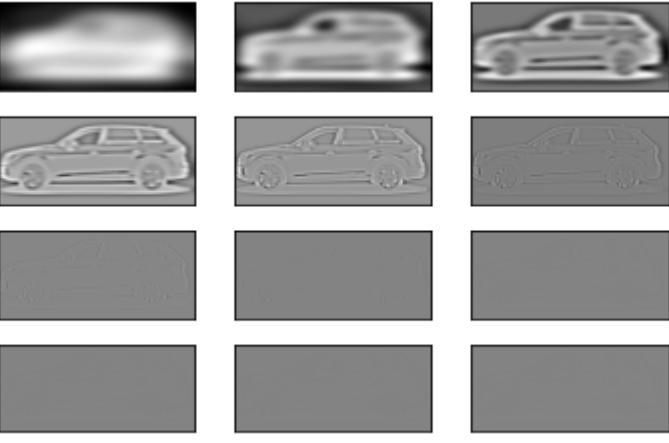
## 6. Photoshop?! Gaussian Filter is There for You!

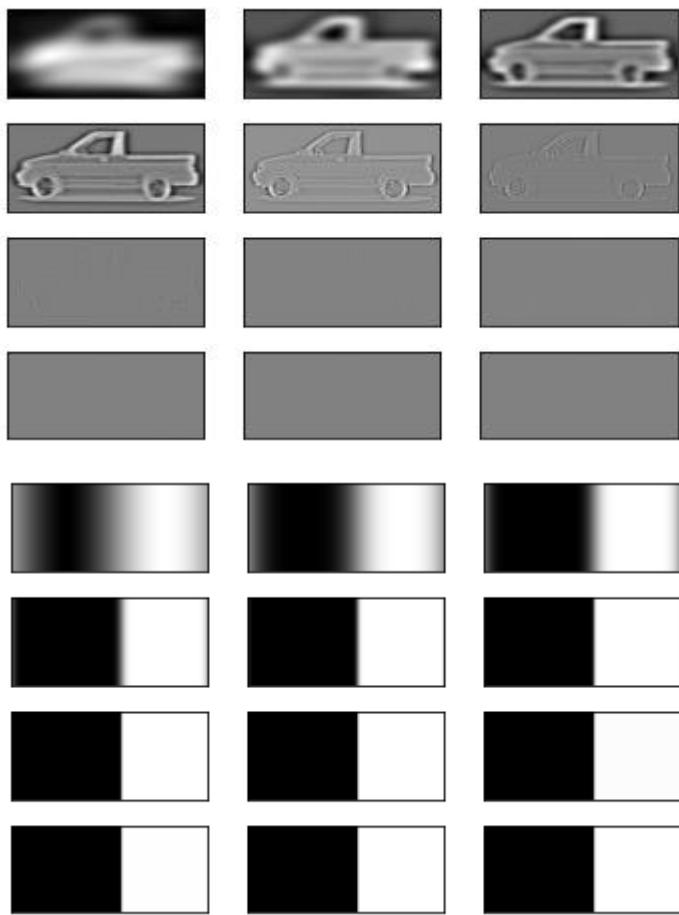
### I. Joint product of Audi and Saipa!

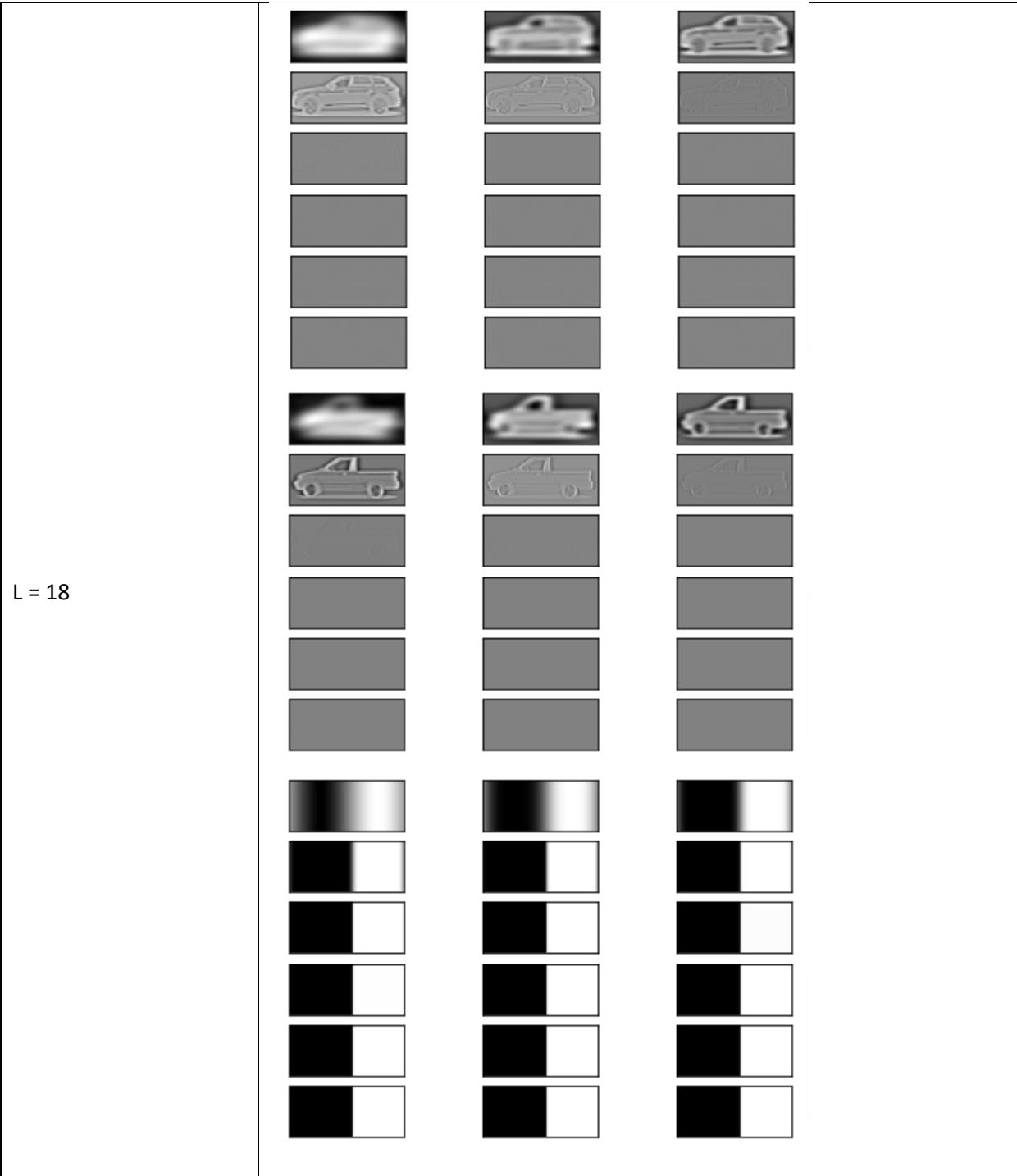
The first task is a simple one; it only contains mixing images in grayscale.

a. Convert the images to grayscale and mix them using the above mentioned method. Your results must be in grayscale. Set different values for the number of levels  $n$ , and report the best result you obtained.



	
$L = 12$	



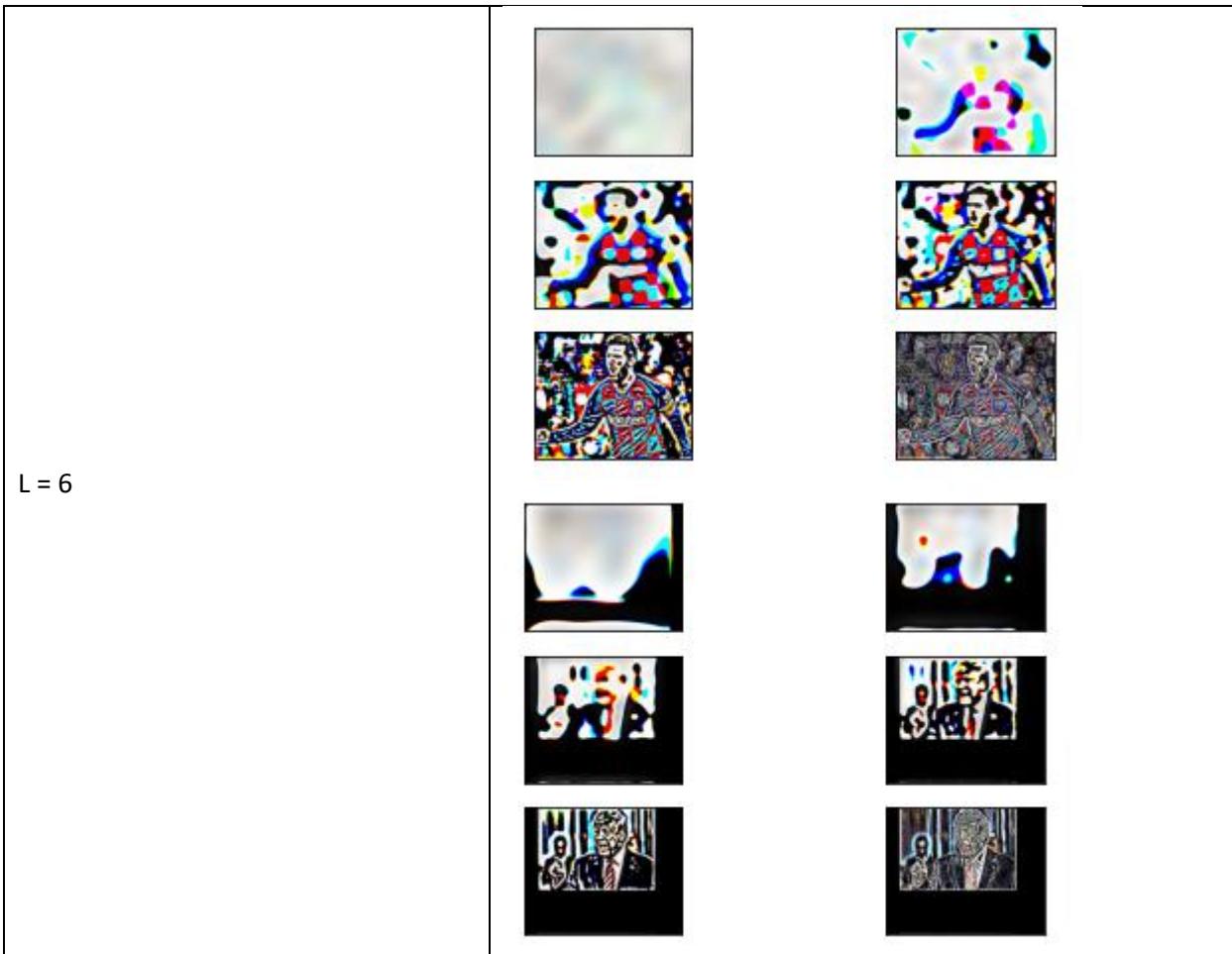


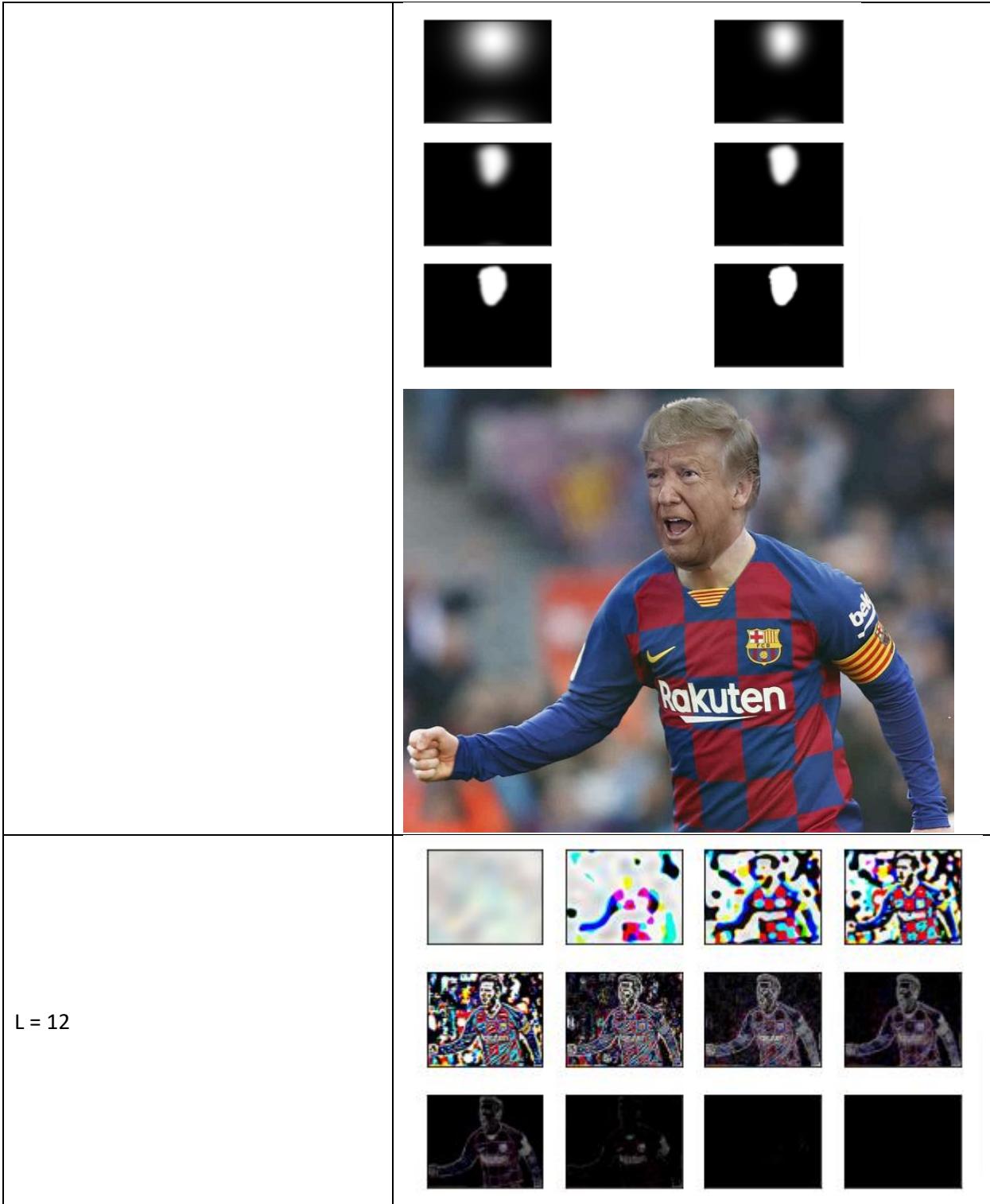


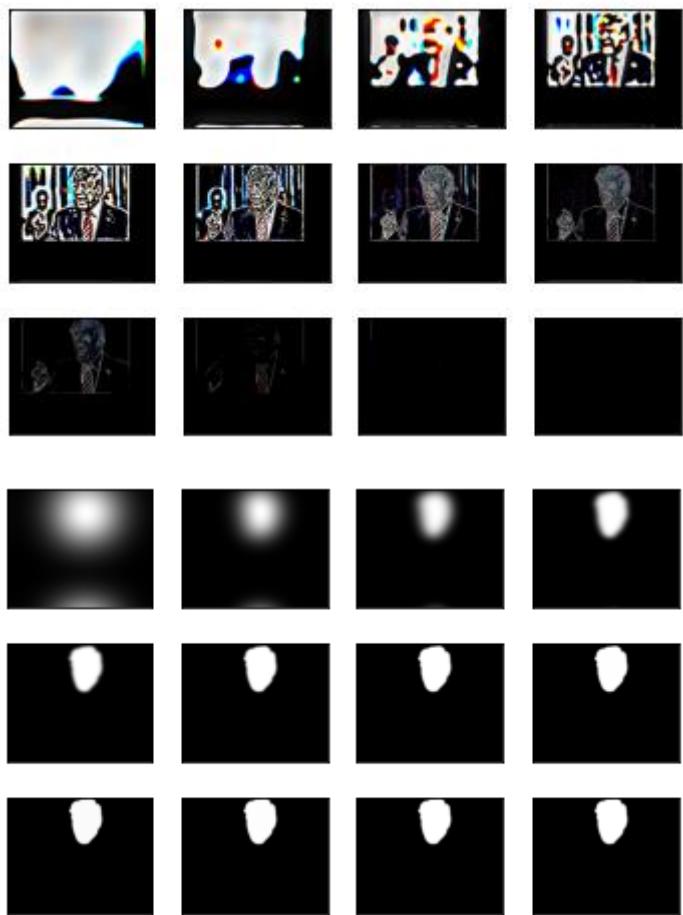
## II. Messi who? We have Donald in our team!

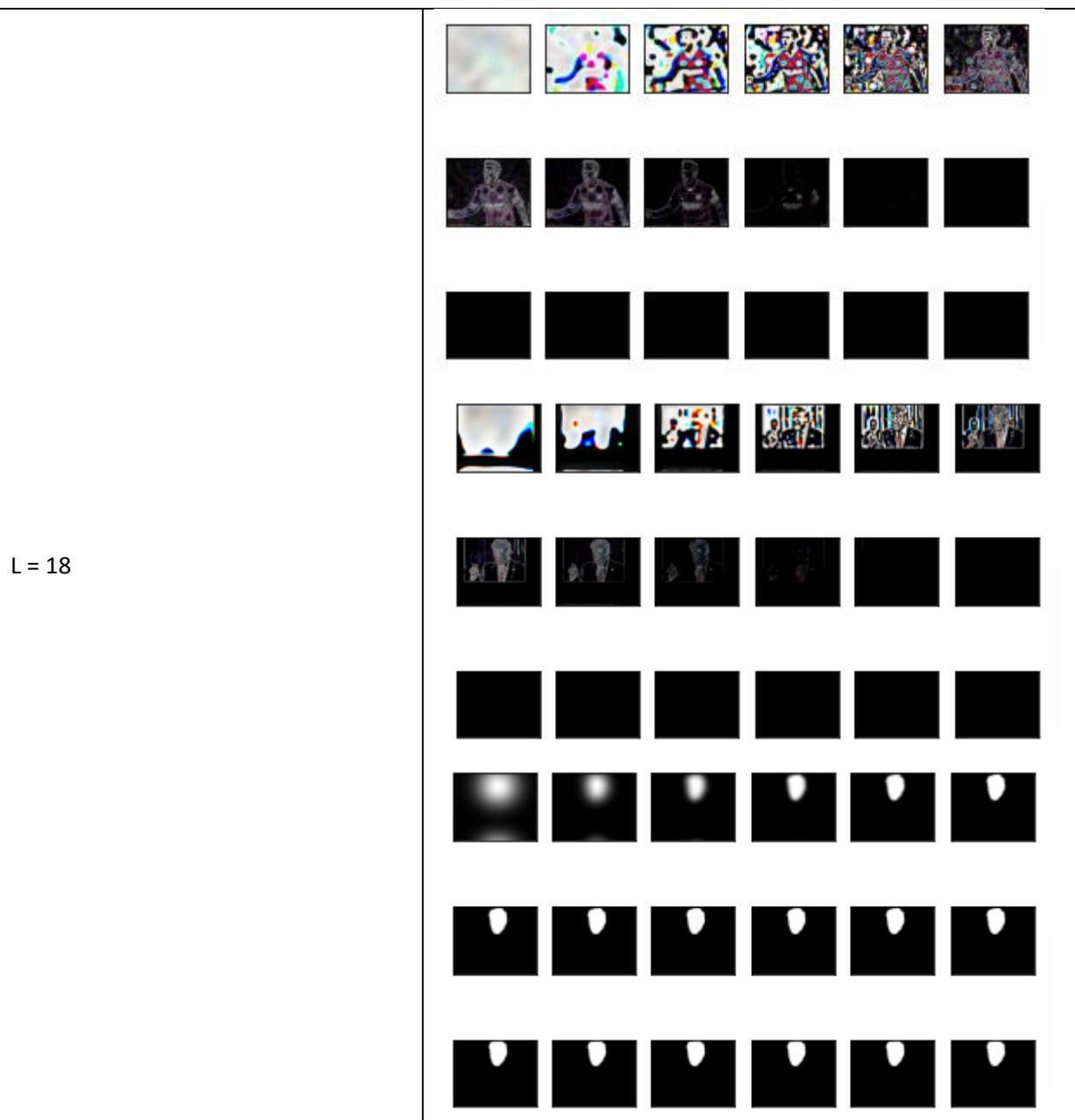
Now a more complicated case. The goal here is to merge images in color space. The above-mentioned method must first be applied to each channels separately, and then the blended results in each channel must be combined together to construct a final image.

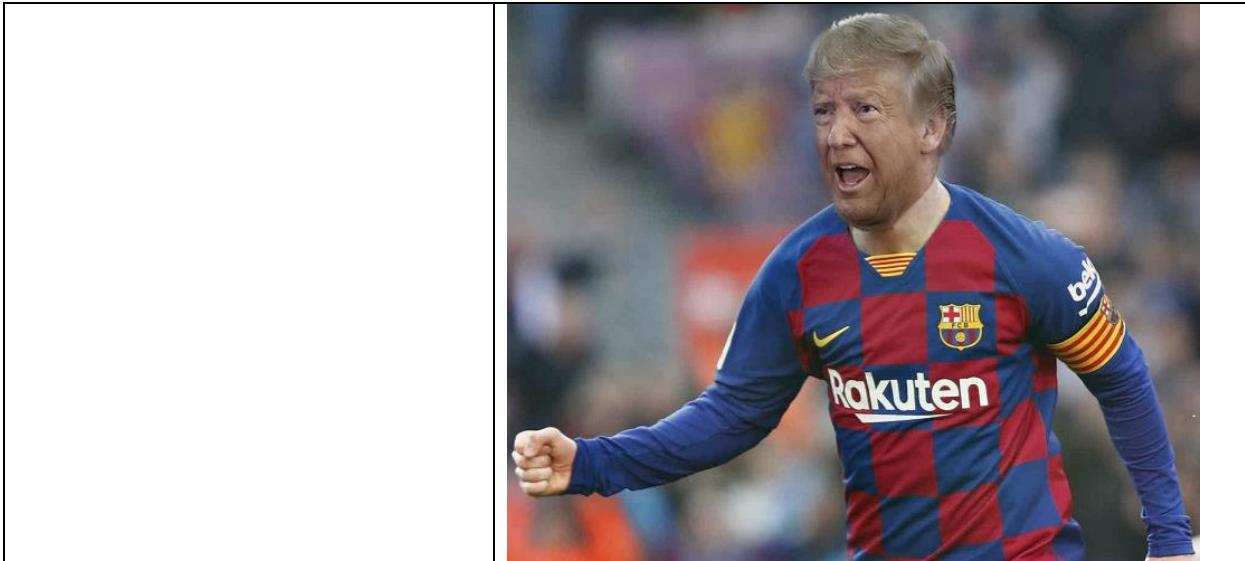
b. Apply the method to different channels of the first image and obtain a final colored result. Set different values for the number of levels  $n$ , and report the best result you obtained.











### III. Just a reminder: Wash your hands!

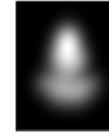
In the third and final task, the goal is to mix images in color space, while the mask is not given.

c. Define a zero matrix of the size of the t-shirt image, and then place the note image in the centre of it. Use this matrix to create a mask (set the nonzero values to 1) and a modified image of the note (set the zero values to the background color) to use in the above-mentioned algorithm.



d. Apply the method to different channels of the first image, and then obtain a final colored result. Set different values for the number of levels  $n$ , and report the best result you obtained.

$L = 6$

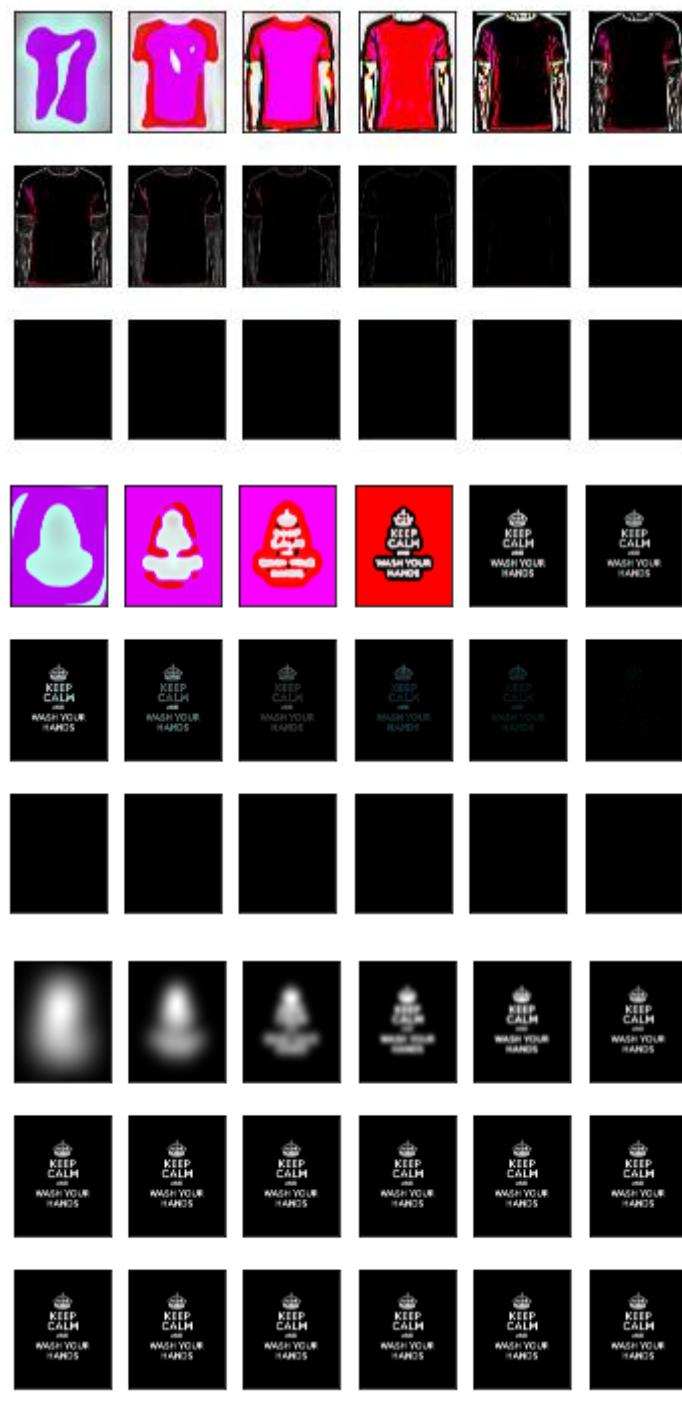




L = 12



$L = 18$





با توجه به تصاویر بالا به نظر می رسد با بیشتر شدن تعداد لایه ها، میزان blending بالاتر می رود و تصویر بهتری حاصل می دهد. البته با مشاهده تصاویر میانی، می توان فهمید که کم کم با بیشتر شدن تعداد لایه ها، تغییر چندانی در ماسک ها و لاپلاسین ها رخ نمی دهد، بنابراین تعداد بیشتر آنها اضافی خواهد بود.

## 7. Some Explanatory Questions

Please answer the following questions as clear as possible:

a. Mathematically explain how rotation and scaling affects phase plot of an image in the Fourier domain.

b. How and under what circumstances the discrete Fourier transform relates to the Fourier transform?

وقتی فرکانس نمونه گیری زیاد باشد و به سمت پی نهایت برود، مشابه حالت پیوسته می شود.

c. In practice, even if Nyquist criterion is satisfied, aliasing cannot be avoided in general. Explain why.

d. The shifted 2D DFT usually forms the shape of plus sign, as it has larger values along the horizontal and vertical axes. Explain why.

e. Although performing convolution in 3D space, for example in video processing, and considering three dimensions of space-space-time, is possible, but it's a poor approach which has almost never been used in video processing. Explain why.