بنامر اهورامزدا

پروژه مبانی برنامه نویسی کامپیوتر

- ۱) قبل از انجام پروژه از نصب بودن پکیج matplotlib اطمینان حاصل کنید. در صورت عدم وجود، به کمک pip میتوانید آن را نصب کنید. pip install matplotlib
- ۲) فایل ارسالی پروژه شما باید یک فایل zip. و یا rar. شامل کد برنامه شما (با پسوند ipynb. و یا py) و یک گزارش pdf. باشد. در گزارش خود، لازم است که در خصوص نحوه عملکرد برنامه به صورت کلی و خط به خط، توضیح دهید. برای تمامی توابع نیز باید توضیح کامل داده شود.
 - ۳) نیازی به کشیدن فلوچارت نیست.
 - ٤) نام فايل zip. و يا rar. ارسالي بايد شامل نام خانوادگي شما باشد.
 - 0) علاوه بر ارسال فایلها، در نهایت ارایه پروژه به صورت شفاهی (در حد چند سوال و جواب) هم انجام خواهد شد.

پردازش متن در تصاویر یکی از کاربردهای حوزه بینایی ماشین است. نیاز به استخراج متن از تصاویر و ویدنوها در صنایع، فعالیت های شهری، فعالیت های تحقیقاتی، و ... به شدت افزایش یافته است. البته در عمده این فعالیت ها به دلیل جامعیت و گسترده بودن نیاز، پردازش تصاویر توسط شبکه های عصبی مصنوعی تحقق می یابد ولی استفاده از این سیستم ها نیازمند سخت افزار قوی و پرسرعت خواهد بود. در این پروژه یک سیستم استخراج کننده متن از تصاویر را بررسی خواهیم کرد که در عین سادگی و بدون استفاده از علوم داده و شبکه های عصبی، میتواند بخشی از نیاز به پردازش متن را پاسخ دهد. یکی از چالش های پیدا کردن و تشخیص حروف در تصاویر، اندازه متغیر آنها است. برای مثال یک صفحه روزنامه را در نظر بگیرید. در یک صفحه روزنامه، متون مختلف در اندازه ها و فونت های مختلف قرار میگیرند. یافتن و تشخیص متون در اسناد از این دست پیچیدگی های بیشتری دارد. در این پروژه، بر اسناد و تصاویری تمرکز میکنیم که حروف، جایگاه و اندازه مشخص و یکسان دارند.

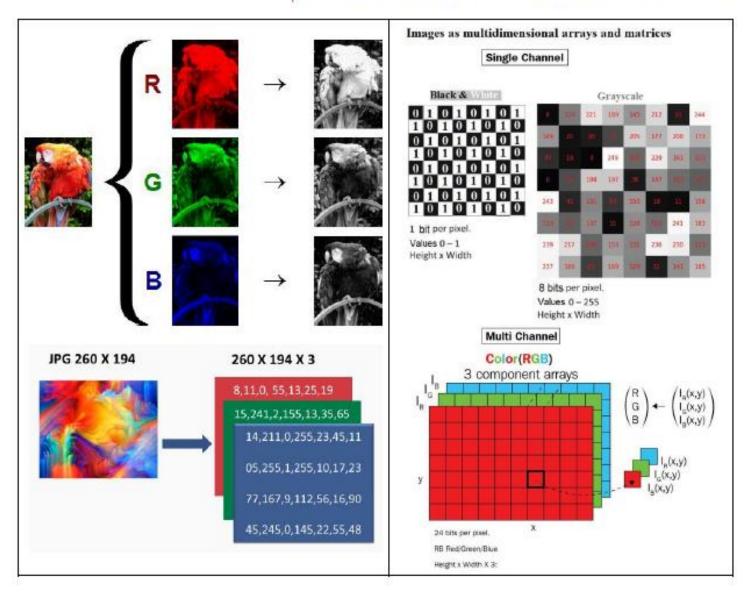
Once upon a time, typing out work or documents used to be quite a laborious and unforgiving experience. Imagine having to press return each time you got to the end of each line! Mistakes in your typing would mean having to start all over again on a fresh piece of paper, or locate some correction fluid, whichwould scar the document for ever!

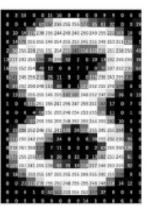
Recreating this experience in the modern world allows you to appreciate how far we have come using technology to hepp produce our really important documents.

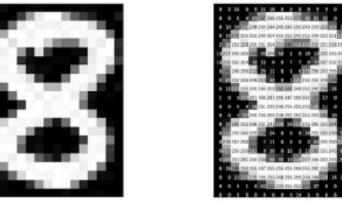
Over on uniquode.com, developers have created an anazing online OverType web programme which recreates the complexities, and unforgiving nature of the typewriter, which can be a great resource for teachers to show pupils how typing used to be, and the many features which we take for granted in our spellchecker and auto-correct technological society.



تصاویر در کامپیوترهای امروزی به صورت آرایه ای از اعداد ذخیره و استفاده میشوند. این اعداد شدتهای نور پیکسلهای تصاویر در لایه های نوری مختلف (برای مثال R و G و B) هستند که در عموم سیستم ها (سیستم های ذخیره ۸بیتی) در بازه [0, 255] مقدار میگیرند. در این پروژه به جهت سادگی با تصاویر grayscale که فقط شامل ۱ لایه رنگی هستند مواجه هستیم.



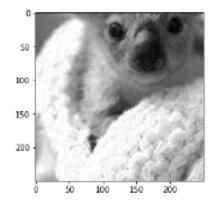




همانطور که گفته شد، تصاویر آرایه ای از اعداد هستند. برای مثال تصویر زیر را در نظر بگیرید که شامل ۱ لایه رنگی، و 800x450 پیکسل در لایه است. به کمک زبان برنامه نویسی Python و پکیج Numpy، فایل این عکس را در یک برنامه بارگزاری کرده و میتوانیم مقادیر هر پیکسل آن را بخوانیم. برای بارگزاری تصاویر و تبدیل آن به آرایه در Numpy از تابع آماده (load_image) استفاده کنید.

```
import utils
                                                 50
 2
                                                100
    img = utils.load_image("img.jpg")
 3
                                                150
 4
                                                200
 5
                                                250
 6
    utils.show image(img) # ======>
                                                300
                                                350
 8
   print(img[0, 0])
 9
                                                400
    print(img[5:10, :10])
10
                                                        100
                                                                                         700
                                                   Ó
                                                             200
                                                                   300
                                                                         400
                                                                              500
                                                                                    600
```

ابتدا تصویر را بارگذاری کرده و ابعاد آن را بدست آورید. سپس یک مربع 250x250 از وسط تصویر slice کنید و نمایش دهید. برای نمایش، از تابع آماده ()show_image استفاده کنید.



همانطور که گفته شد، در پروژه استخراج متن، حروف جایگاه و اندازه مشخصی دارند. تصویر text.png را توسط تابع آماده بارگزاری کرده و نمایش دهید. همانطور که مشاهده میکنید، این تصویر مشابه یک grid از حروف با سایز و جایگاه ثابت است. هدف از انجام این پروژه استخراج این حروف و ذخیره متن آن است.

تصویر index.png شامل تصویر میانگینی از ۲۱ حرف زبان انگلیسی است. حروف به ترتیب A, B, ... از بالا به پایین، با اندازه مشخص (برابر اندازه حروف تصویر متنی) قرار گرفته اند. به کمک تابع از پیش نوشته شده ()load_image و به کمک slicing در Numpy، میتوانید تصاویر معیار حروف انگلیسی را مشاهده و بررسی نمایید.

```
WHAT IS COMPUTATION

THIS HAS ALWAYS BEEN THE MOST FUNDAMENTAL QUESTION OF OUR FIELD

IN THE THIRTIES AS THE FIELD WAS STARTING THE ANSWER WAS THAT COMPUTATION WAS THE ACTION OF PEOPLE WHO OPERATED CALCULATOR MACHINES

BY THE LATE FOUNTIES. THE ANSWER WAS THAT COMPUTATION WAS STEPS CARRIED OUT BY AUTOMATED COMPUTERS TO PRODUCE DEFINITE OUTPUTS

THAT DEFINITION DID VERY WELL IT REMAINED THE STANDARD FOR NEARLY FIFTY YEARS

BUT IT IS NOW BEING CHALLENGED PEOPLE IN MANY PIELDS HAVE ACCEPTED THAT COMPUTATIONAL THINKING

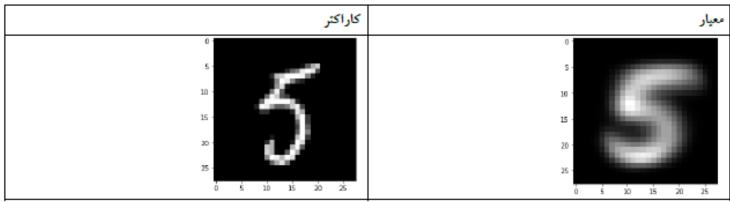
IS A WAY OF APPROACHING SCIENCE AND ENGINEERING

THE INTERNET IS FULL OF SERVERS THAT PROVIDE NONSTOP COMPUTATION ENDLESSLY

RESEARCHERS IN BIOLOGY AND PHYSICS HAVE CLAIMED THE DISCOVERY
OF NATURAL COMPUTATIONAL PROCESSES THAT HAVE NOTHING TO DO WITH COMPUTERS

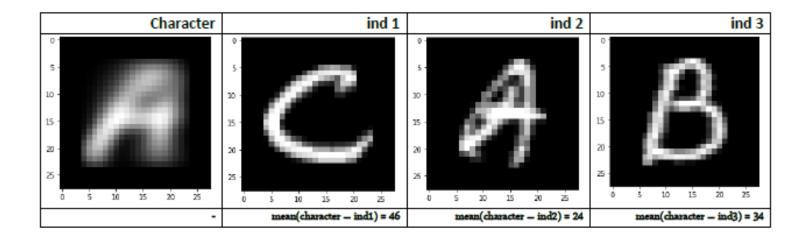
NOW MUST OUR DEFINITION EVOLVE TO ANSWER THE CHALLENGES OF BRAINS COMPUTING ALGORITHMS NEVER TERMINATING BY DESIGN COMPUTATION AS A NATURAL OCCURRENCE AND COMPUTATION WITHOUT GOMPUTERS
```

برای استخراج حروف از متن، ابتدا (به ترتیب خواندن از بالا-چپ) تصویر هر کاراکتر slice میشود. سپس برای تشخیص حرف، میانگین اختلاف مقادیر پیکسل های آن، با مقادیر پیکسل های تصویر هر حرف معیار گرفته میشود. کمترین میانگین اختلاف بین تصویر کاراکتر و یک حرف معیار، برابری آن کاراکتر با آن حرف معیار را مشخص میکند.



راهنمایی: برای محاسبه میانگین اختلاف، تصویر را به شکل آرایه سطری در آورده و از numpy.mean در محور مناسب کمک بگیرید.

برای مثال، میانگین اختلاف تصویر کاراکتر زیر با تصاویر چند حرف معیار نمایش داده شده اند. مشخصا و با بررسی اعداد، کاراکتر موجود در تصویر، حرف A را نشان میدهد.



تصویر متن دار text.png را در نظر بگیرید. به ترتیب (از بالا-چپ) حروف این تصویر را استخراج کرده و در یک متغیر string ذخیره کنید. این متن را چاپ کرده و سپس در یک فایل متنی به فرمت txt نوشته و ذخیره کنید.

بخش امتیازی: اگر به خروجی الگوریتم توجه کنید، متوجه می شوید که حروف r و k به درستی تشخیص داده نمیشوند. با بررسی تصاویر و الگوریتم و به کمک عملیات ماتریسی توضیح دهید چگونه میتوان این خطا در تشخیص را از بین برد. راه حل خود را پیاده سازی کرده و دوباره خروجی بگیرید.

در یناه اهورامزدا شاد و موفق باشید