

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «КИЄВО-МОГИЛЯНСЬКА АКАДЕМІЯ»

Факультет інформатики

Кафедра математики

ЛАБОРАТОРНА РОБОТА №2

з курсу «Математичні методи машинного навчання»

Варіант 2

Виконала:

студентка 3 р.н.

БП «Прикладна математика»

Гак Софія Володимирівна

Київ-2020

Мета:

Провести декомпозицію зеленого каналу вихідних зображень методом головних компонент.

Змодельовати зелений канал заданих зображень за допомогою марківських ланцюгів першого порядку.

Частка виконаної роботи:

Повністю виконана (п. 1-4).

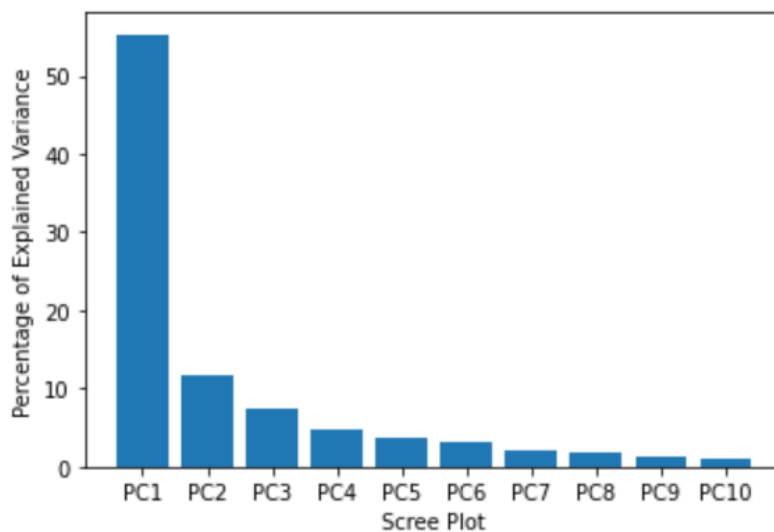
1) Сформувати тестову вибірку зображень з вихідного пакету;

Вибірку було сформовано аналогічно до першої лабораторної роботи.

2) Провести декомпозицію каналу зеленого кольору тестових зображень з використанням методу головних компонент (PCA):

Для реалізації методу головних компонент застосувала бібліотеку `sklearn`. у функції нижче створюється об'єкт `PCA()`, в якому важливим аргументом є `n_components`, що позначає відсоток або кількість головних компонент, які хочемо отримати (якщо `None`, тоді з документації `n_components == min(n_samples, n_features) - 1`, тобто в нашому випадку, декомпозиючи картинку, за кількість компонент вибирається або її ширина, або висота - залежно від того, що менше). Саму ж сингулярну декомпозицію забезпечує метод `fit_transform()`, який знаходить сингулярні вектори (вони ж - головні компоненти), відповідні їм сингулярні числа, дисперсія значень яскравості пікселів зеленого каналу і т.п. Для реконструкції зображення застосовується метод `inverse_transform`.

Частка поясненої дисперсії (%) кожної з компонент (виведені перші 10)



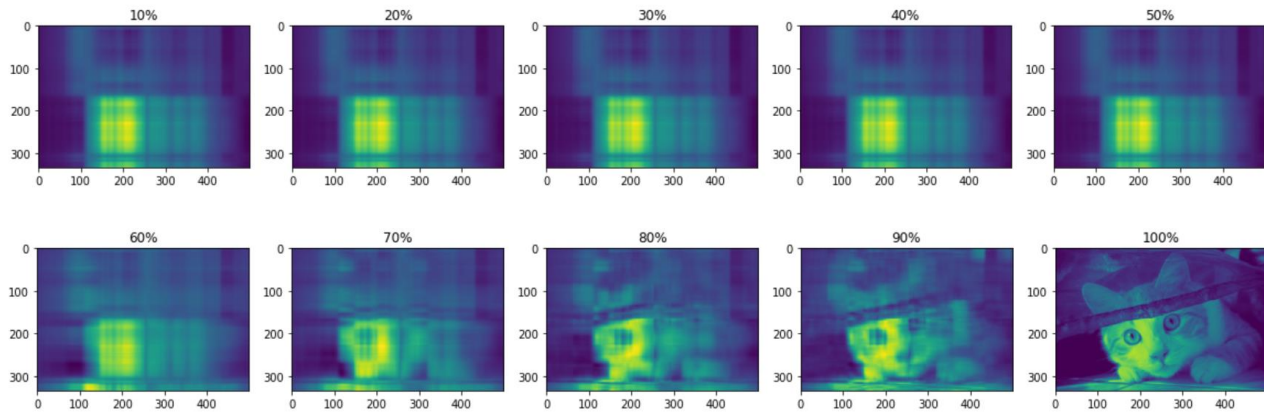
- а) Впорядкувати отримані компоненти вихідного зображення в порядку зменшення значень сингулярних чисел (від найбільшого S_{max} до найменшого S_{min} значення);

За замовченням сингулярні числа вже відсортовані в порядку зменшення.

Найбільше з них - $1.32332115e+04$.

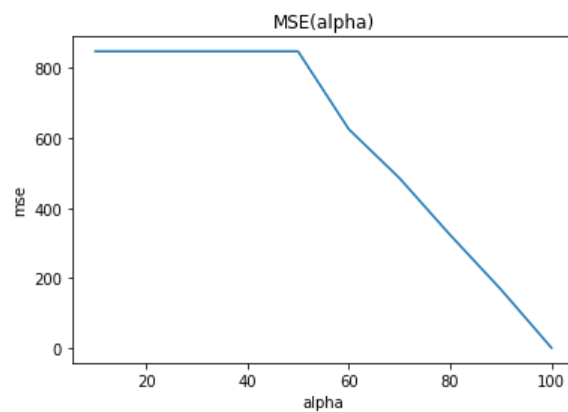
- b) Провести реконструкцію зображення при використанні лише частки ($\alpha\%$) компонентів розкладу, що характеризуються відмінними від нуля сингулярними числами ($s_i \neq 0$). Розглянути випадок, коли α змінюється від $\alpha_{min} = 10\%$ до $\alpha_{max} = 100\%$ з кроком $\Delta\alpha = 10\%$.

Візуальні результати реконструкції одного з тестових зображень залежно від частки використаних головних компонент.



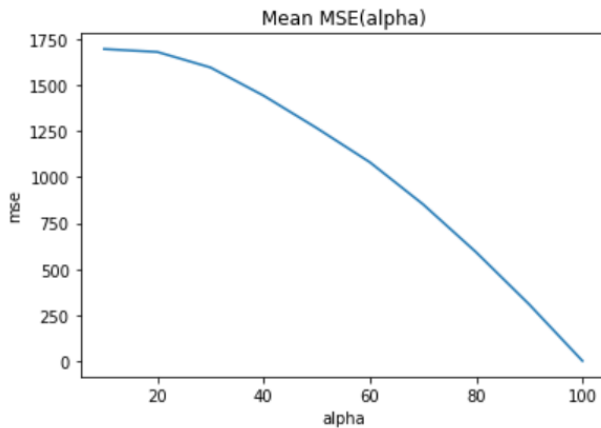
- c) Порівняти вихідне та реконструйоване зображення за показником середньоквадратичного відхилення (MSE).

	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
MSE	848.031981	848.031981	848.031981	848.031981	848.031981	625.480141	485.577072	323.019105	167.433249	2.169044e-26



- d) Побудувати графіки залежності $\overline{MSE}(\alpha)$, де \overline{MSE} – значення середньоквадратичного відхилення між вихідним та реконструйованим зображеннями, усереднені по тестовому пакету;

	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
Mean MSE	1697.948712	1682.171739	1597.620534	1444.325154	1267.735455	1081.515872	852.842967	591.520586	307.047362	4.798322e-26



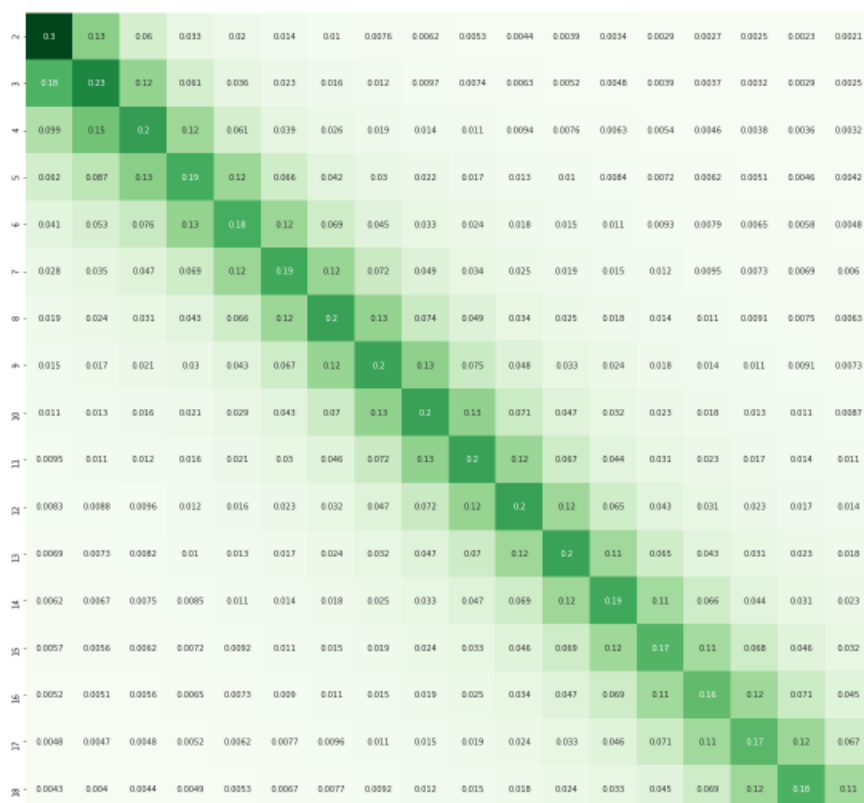
3) Провести моделювання каналу зеленого кольору тестових зображень з використанням марківських ланцюгів першого порядку M_1 :

- a) Отримати стохастичну матрицю для каналу зеленого кольору при обробці пікселів (згідно номеру студента в списку групи, за модулем кількості варіантів):

- i) По горизонталі, зліва направо – $M_1^{\rightarrow}(I_{x,y}, I_{x+1,y})$;
- ii) По горизонталі, справа наліво – $M_1^{\leftarrow}(I_{x,y}, I_{x-1,y})$;
- iii) По вертикалі, зверху вниз – $M_1^{\downarrow}(I_{x,y}, I_{x,y+1})$;
- iv) По вертикалі, знизу вгору – $M_1^{\uparrow}(I_{x,y}, I_{x,y-1})$;
- v) По головній діагоналі – $M_1^{\searrow}(I_{x,y}, I_{x+1,y+1})$;
- vi) По головній діагоналі – $M_1^{\swarrow}(I_{x,y}, I_{x-1,y-1})$;
- vii) По побічній діагоналі – $M_1^{\swarrow}(I_{x,y}, I_{x-1,y+1})$;
- viii) По побічній діагоналі – $M_1^{\searrow}(I_{x,y}, I_{x+1,y-1})$;

- b) В протоколі роботи графічно показати вид марківського ланцюга для діапазону яскравості пікселів $I_{x,y} \in [i; i \times 10]$, де i – номеру студента в списку групи;

Із матриці переходу Q було відібрано діапазон таких станів - [2;20]. І цю підматрицю було відображено у вихідному коді лабораторної у вигляді «кольорової» матриці, кожна клітинка якої має свою інтенсивність кольору в залежності від ймовірності переходу, що лежить в ній. Що вища інтенсивність, то вища ймовірність.



- c) Для отриманих марківських ланцюгів перевірити виконання властивості регулярності та рекурентності після проходження 5 ітерацій роботи.

Марківський ланцюг є *регулярним*, якщо його матриця переходів є регулярною. Матриця переходів є регулярною, якщо деякий її степінь містить лише додатні елементи.

Марківський ланцюг є *рекурентним*, якщо кожен його стан рекурентний, тобто вийшовши з деякого стану, ми колись до нього повернемося. Це можна перевірити за матрицею переходу: на певному кроці (в нашому випадку - на 5) її діагональні елементи мають бути ненульовими.

Перевірка побудованого ланцюга показала, що він володіє властивостями рекурентності і регулярності.