Міністерство освіти і науки України

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «КИЄВО-МОГИЛЯНСЬКА АКАДЕМІЯ»

Факультет інформатики



**Протокол до лабораторної роботи №2**

**З дисципліни «Математичні методи машинного навчання»**

Виконав

студент 4 курсу

факультету інформатики

Миколайчик Я. А.

1. Завдання лабораторної роботи

1. Сформувати тестову вибірку зображень з вихідного пакету;
2. Провести декомпозицію каналу зеленого кольору тестових зображень з використанням методу головних компонент (PCA);
3. Для каналу зеленого кольору тестових зображень обчислити наступні характеристики розподілу значень яскравості пікселів:
   1. Впорядкувати отримані компоненти вихідного зображення в порядку зменшення значень сингулярних чисел (від найбільшого до найменшого значення);
   2. Провести реконструкцію зображення при використанні лише частки () компонентів розкладу, що характеризуються відмінними від нуля сингулярними числами (). Розглянути випадок, коли змінюється від до з кроком .
   3. Порівняти вихідне та реконструйоване зображення за показником середньоквадратичного відхилення (MSE).
   4. Побудувати графіки залежності , де значення середньо-квадратичного відхилення між вихідним та реконструйованим зображеннями, усереднені по тестовому пакету;
4. Провести моделювання каналу зеленого кольору тестових зображень з використанням марківських ланцюгів першого порядку :
   1. Отримати стохастичну матрицю для каналу зеленого кольору при обробці пікселів (згідно номеру студента в списку групи, за модулем кількості варіантів):
      1. По горизонталі, зліва направо – ;
      2. По горизонталі, справа наліво – ;
      3. По вертикалі, зверху вниз – ;
      4. По вертикалі, знизу вгору – ;
      5. По головній діагоналі – ;
      6. По головній діагоналі – ;
      7. По побічній діагоналі – ;
      8. По побічній діагоналі – ;
   2. В протоколі роботи графічно показати вид марківського ланцюга для діапазону яскравості пікселів , де номеру студента в списку групи;
   3. Для отриманих марківських ланцюгів перевірити виконання властивості регулярності та рекурентності після проходження 5 ітерацій роботи.

2. Порядок виконання роботи та отримані результати

Робота була виконана на мові Python в форматі Jupyter Notebook та за допомогою середовища JupyterLab.

1. **Формування тестової вибірки зображень з вихідного пакету.**

Для формування вибірки було використано модуль random з заданим відповідним діапазоном значень. У якості зображення, що використовуватиметься як приклад, на якому буде демонструватись робота програми, було вибране перше зображення з масиву випадково створеної вибірки зображень.

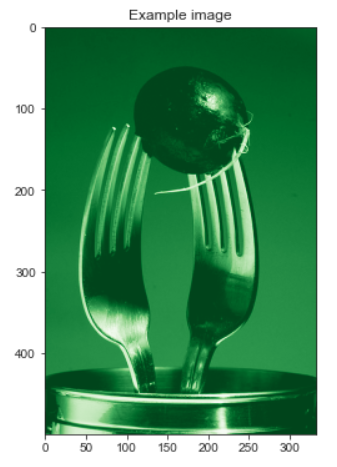


Рис. 1. Зображення для прикладу

1. **Проведення декомпозиції каналу зеленого кольору тестових зображень з використанням методу головних компонент (PCA).**

Для декомпозиції зображень з використанням методу головних компонент був використаний клас PCA модуля sklearn.decomposition. Для демонстрації роботи класу зображення було реконструйовано з використанням 10 компонент розкладу:

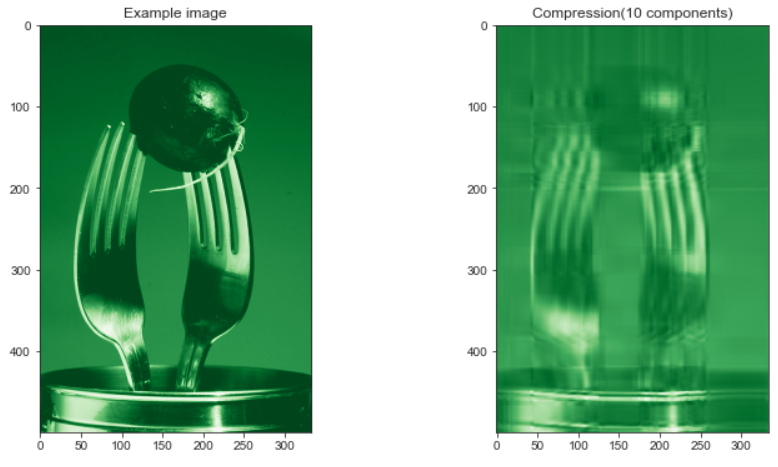
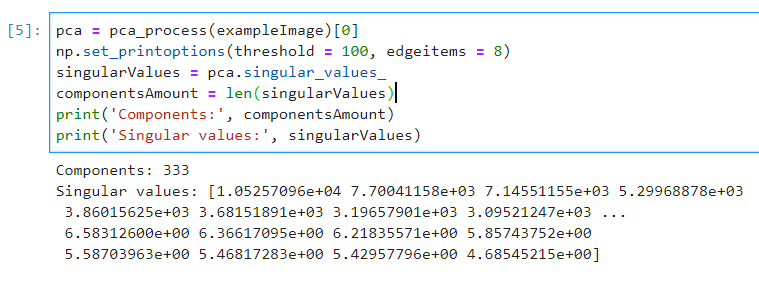


Рис. 2. Приклад реконструкції зображення з використанням PCA

1. **Обчислення характеристик розподілу значень яскравості пікселів для каналу зеленого кольору тестових зображень.**
   1. *Впорядкування отриманих компонент вихідного зображення в порядку зменшення значень сингулярних чисел (від найбільшого до найменшого значення).*



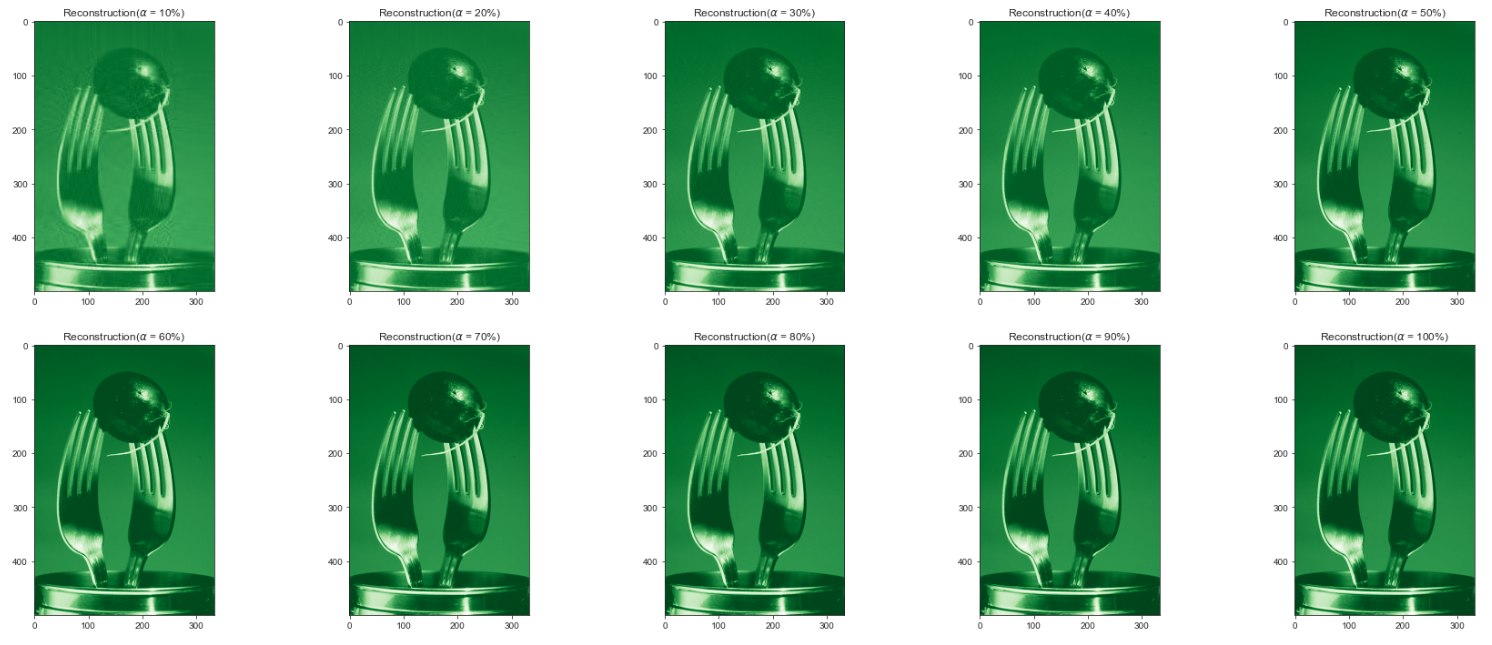
* 1. *Проведення реконструкції зображення при використанні* *лише частки () компонентів розкладу.*

Використання лише частки () компонентів розкладу, що характеризуються відмінними від нуля сингулярними числами ():



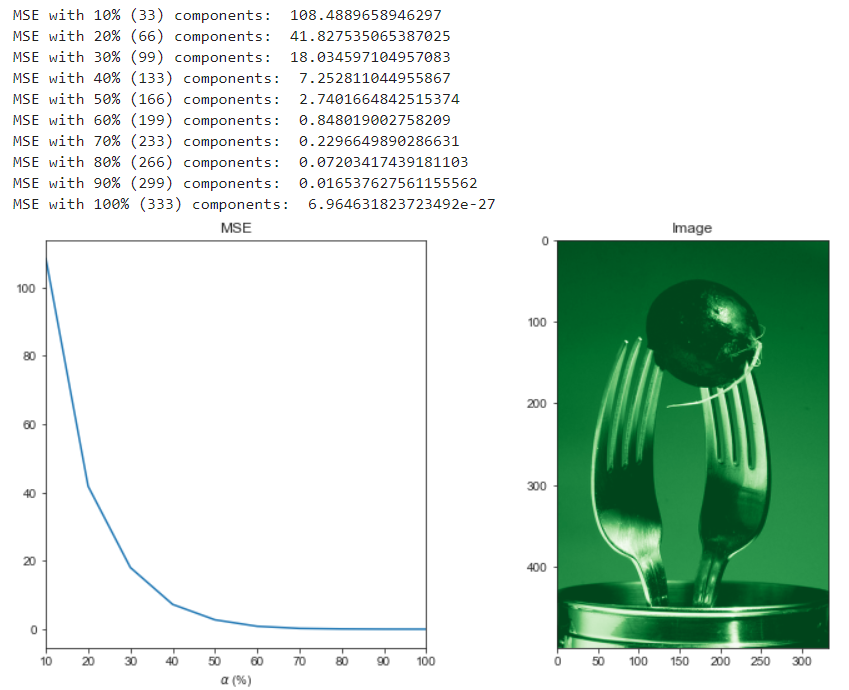
У цьому випадку ненульовими були всі сингулярні числа, тому фактично отримали зображення, ідентичне оригінальному.

Також розглянули випадок, коли змінюється від до з кроком :



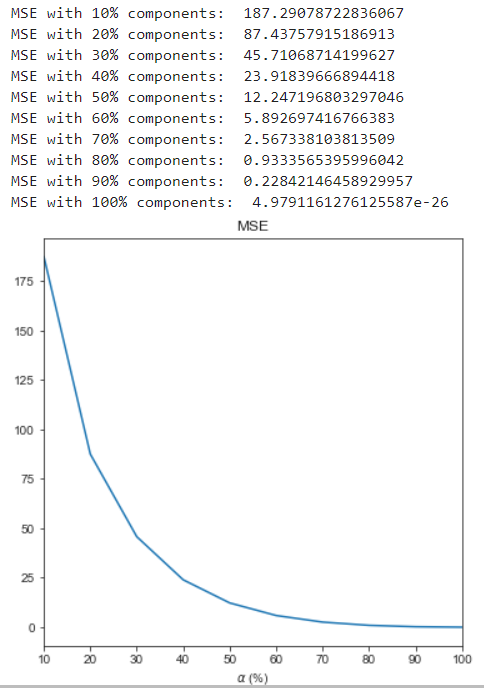
Бачимо, що зі збільшенням кількості використаних компонент при реконструкції зображення вийде чіткішим.

* 1. *Порівняти вихідне та реконструйоване зображення за показником середньоквадратичного відхилення (MSE).*



Бачимо, що чим більша частка компонент використовується в реконструкції, тим менше значення пікселів будуть відрізнятись між реконструйованим і вихідним зображенням. Для наведеного зображення вже починаючи з 60% компонент втрати в якості стають мізерними і непомітними (<1%).

* 1. *Побудувати графіки залежності , де значення середньо-квадратичного відхилення між вихідним та реконструйованим зображеннями, усереднені по тестовому пакету.*

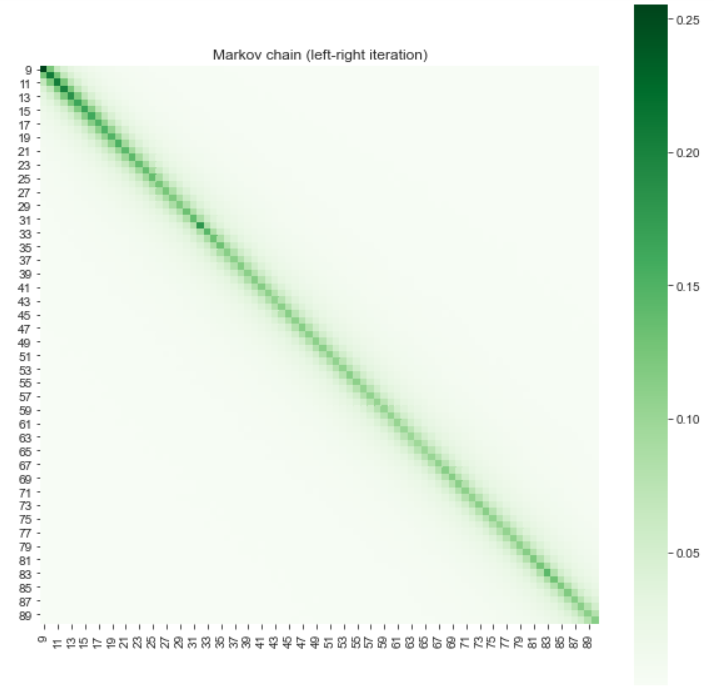


А от опрацювавши всю вибірку, бачимо, що середньоквадратичні втрати <1% досягаються при використанні в реконструкції як мінімум 80% компонент розкладу. Тобто бачимо, що оптимальна кількість використаних компонент різниться в залежності від зображення, але 80% швидше за все буде достатньо для втрати в якості <1%.

1. **Провести моделювання каналу зеленого кольору тестових зображень з використанням марківських ланцюгів першого порядку .**

Побудова стохастичної матриці для каналу зеленого кольору у моєму варіанті відбувалась при обробці пікселів по горизонталі, зліва направо.

Графічне зображення виду марківського ланцюга для діапазону яскравості пікселів , де i - номеру студента в списку групи (9), відбувається у вигляді heatmap(теплової карти):



З графіку бачимо, що значення яскравостей сусідніх пікселів з найбільшою ймовірністю переходять в такі самі, або такі, які відрізняються лише на кілька одиниць. Це означає, що різкі переходи яскравості зеленого кольору між пікселями у зображеннях зустрічаються дуже рідко, а зазвичай відбуваються плавно.

Також після цього для отриманого марківського ланцюга було перевірено виконання властивостей регулярності та рекурентності після проходження 5 ітерацій роботи. Обидві властивості виконуються.

3. Висновки

В ході лабораторної роботи ми сформували вибірку зображень, провели декомпозицію каналу зеленого кольору та обчислили її характеристики. Показали обернену залежність між кількістю компонентів розкладу, використаних в реконструкції, та відхиленням від оригінального зображення.

Провели моделювання каналу зеленого кольору тестових зображень з використанням марківських ланцюгів першого порядку, побудувавши стохастичну матрицю для каналу зеленого кольору при обробці пікселів по горизонталі, зліва направо.

Графічно показали вид марківського ланцюга для діапазону яскравості пікселів від 9 до 90 та побачили частотну перевагу плавного переходу значень яскравості над різким.

Підтвердили виконання властивостей регулярності та рекурентності марківського ланцюга після проходження 5 ітерацій роботи.