

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
Іванюк-Скульський Б. В.

1. Завдання лабораторної роботи

- 1) Сформулювати 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:
 - a) Впорядкувати отримані компоненти вихідного зображення в порядку зменшення значень сингулярних чисел (від найбільшого s_{max} до найменшого s_{min} значення);
 - 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 $\Delta\alpha = 10\%$.
 - c) Порівняти вихідне та реконструйоване зображення за показником середньоквадратичного відхилення (MSE).
 - d) Побудувати графіки залежності $\overline{MSE}(\alpha)$, де \overline{MSE} – значення середньоквадратичного відхилення між вихідним та реконструйованим зображеннями, усереднені по тестовому пакету;
- 4) 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 – $M_1^{\prime}(I_{x,y}, I_{x-1,y+1})$;
 - viii) По побічній діагоналі – $M_1^{\wedge}(I_{x,y}, I_{x+1,y-1})$;
 - 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;
 - c) 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 Jupyter.

- 1) Для формування вибірки було використано модуль random відповідним діапазоном значень.
- 2) Декомпозиція зображення за допомогою PCA

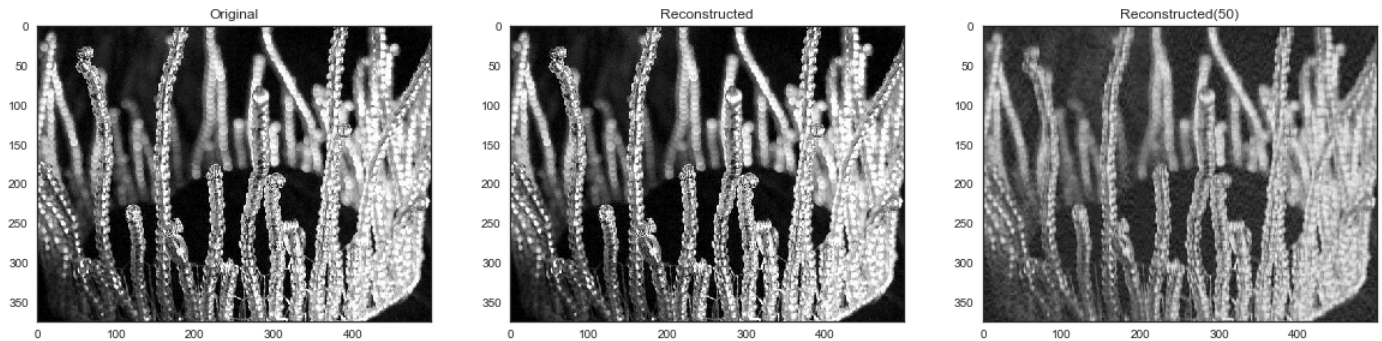


Рис.1. Оригінальне зображення, відновлене на всіх компонентах, відновлене на 50 компонентах

Бачимо, що при збільшенні компонент зображення відновлюється більш точно.

- 3) Після впорядкування отриманих компонент зображення провів реконструкцію зображення від 10% компонент зображення до 100%.

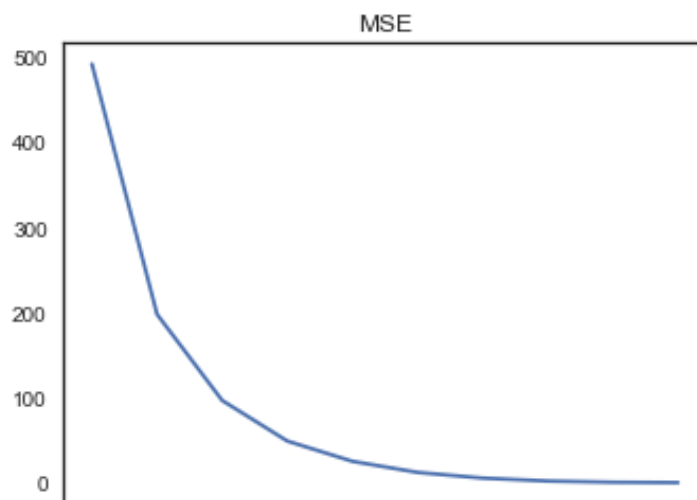


Рис.2. Значення похибки при відновленні зображення з відповідним відсотком компонент

4) MSE похибка для усіх зображень вибірки зображена на графіку нижче

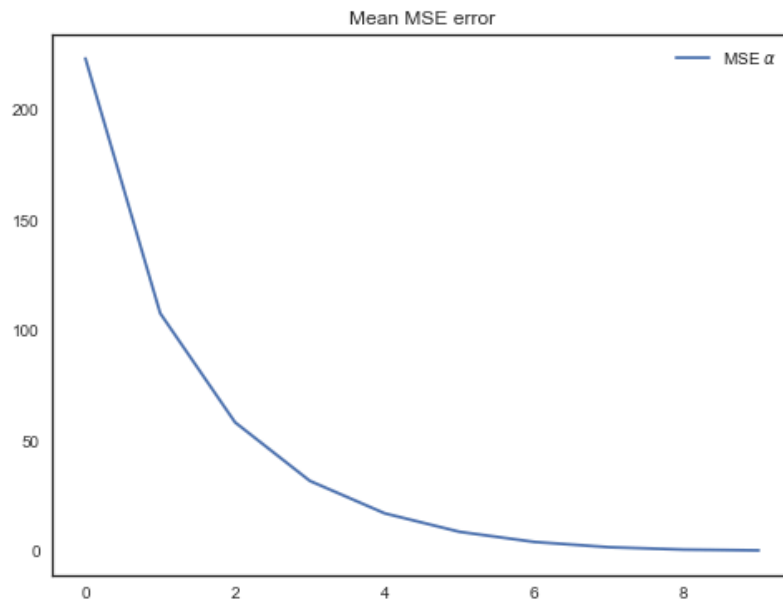
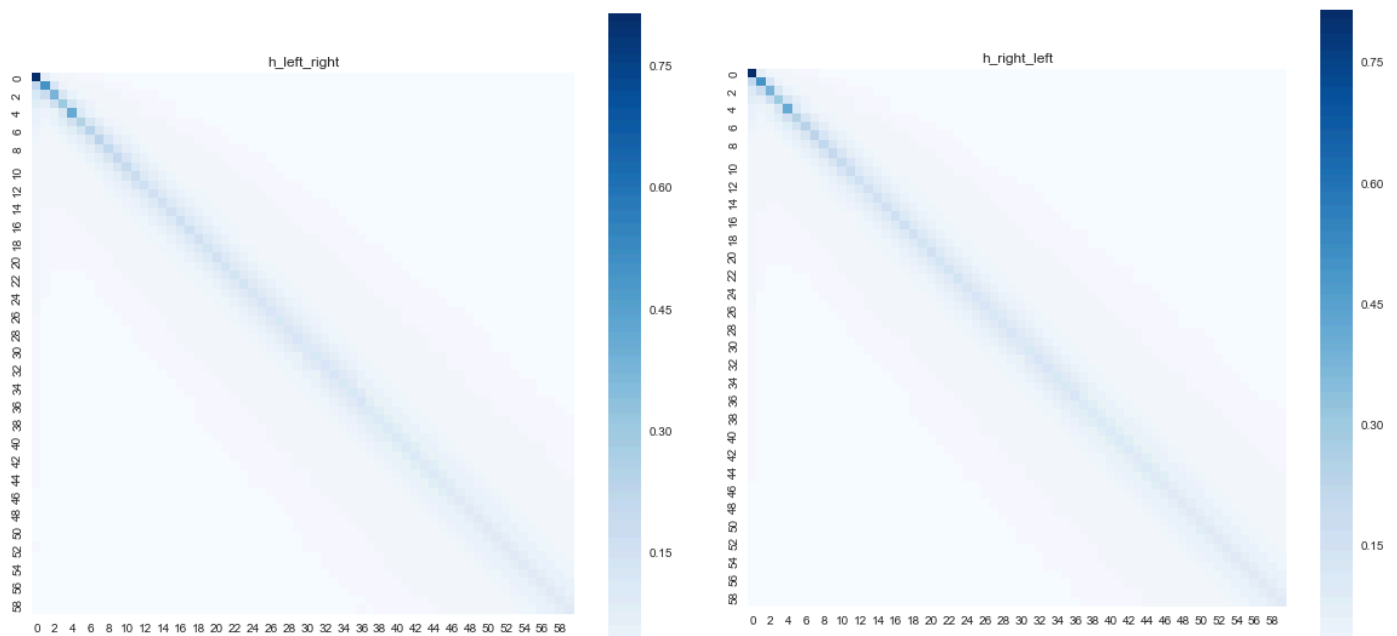
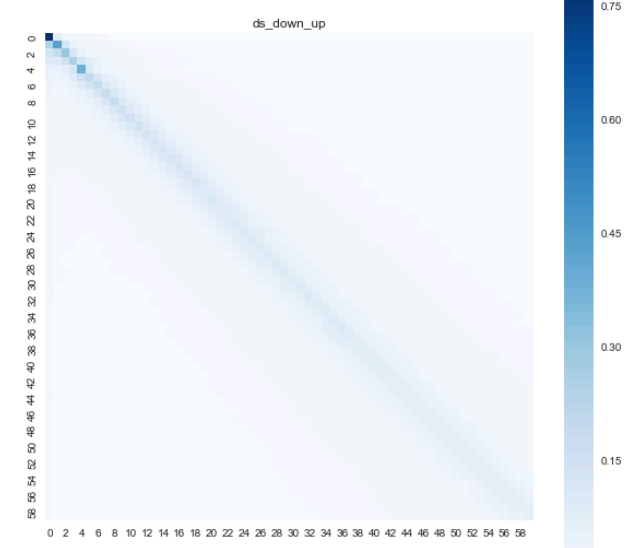
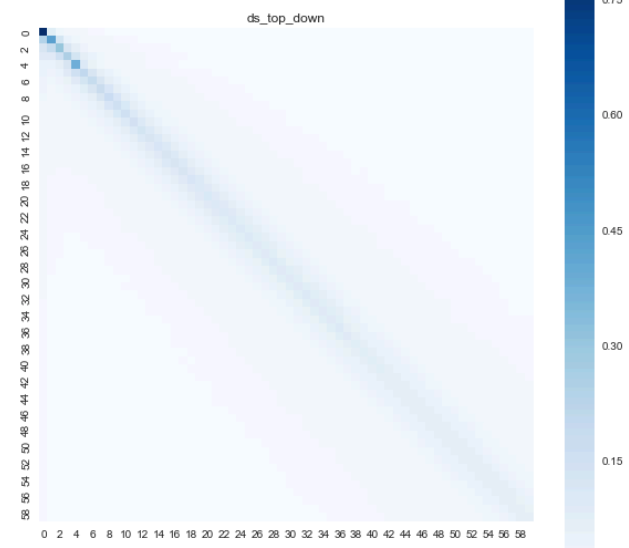
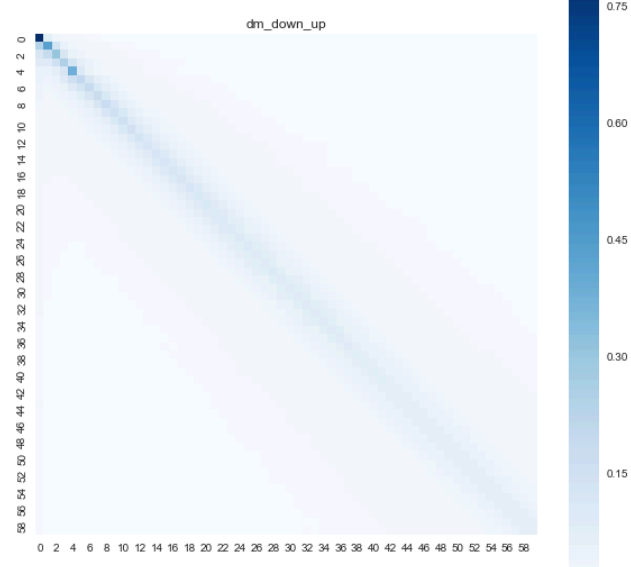
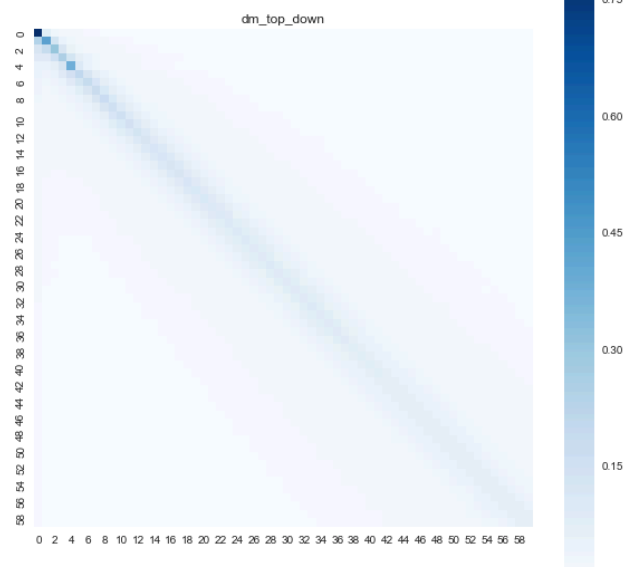
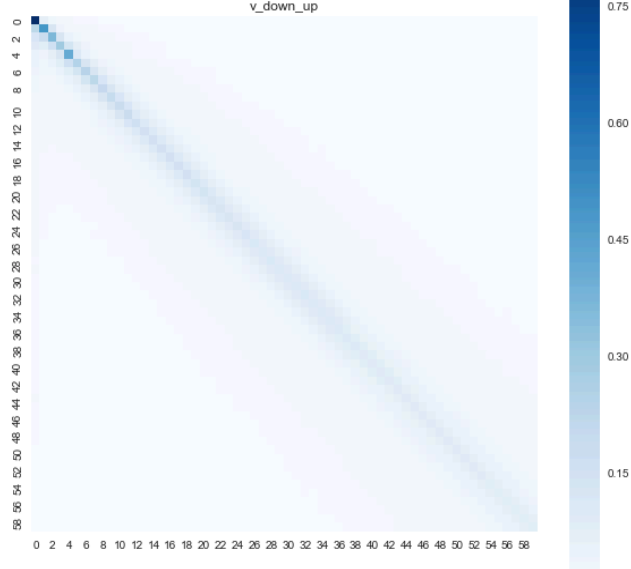
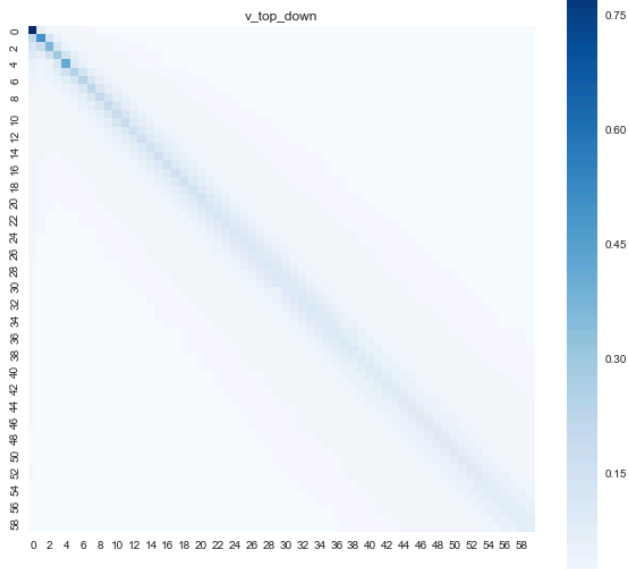


Рис.3. Загальне значення похибки при відновленні зображення

Як і при відновленні одного зображення, чим менше компонент ми використовуємо – тим гірше похибка.

5) Після моделювання зображень марківськими ланцюгами першого порядку, я отримав наступні графіки ланцюгів.





З графіків вище можна бачити, що досить часто сусідні значення пікселів переходять самі в себе. Якщо розглядати достатньо маленьку вибірку зображень, то можна бачити іншу картину – розподіл ймовірностей перейти в інші значення пікселів. В графіку загальної вибірки це можна пояснити тим, що різкі переходи між сусідніми пікселями зглажуються у зальній вибірці.

- 6) Після проходження 5 ітерацій роботи отриманих марківських ланцюгів перевірів на рекурсивність на регулярність. Усі ланцюги є рекурентними та регулярними

3. Висновки

З отриманих даних MIRFlickr відібрав випадковим чином 250 картинок. Відібрав зелений канал та провів декомпозицію зображення за допомогою методу PCA. Подивився на MSE у відновленні зображень з різною кількістю компонент (від 10% до 100%). Потім змодельював канал зеленого кольору за допомогою ланцюгів маркова: по горизонталі зліва направо, по горизонталі справа наліво, по вертикалі згори вниз, по вертикалі знизу вгору, по головній діагоналі зверху вниз, по головній діагоналі знизу вгору, по побічній діагоналі зверху вниз, по побічній діагоналі знизу вгору. Візуалізував матриці переходів та перевірів їх на рекурентність та регулярність.