

Metoda bag of features (worka cech)

Ewa Kurek

May 26, 2021

Pierwszy etap worka słów to wyodrębnienie wszystkich słów z wszystkich tekstów. Te słowa tworzą nasz słownik

- film
- jest
- bardzo
- przerażający
- i
- długi
- pouczający
- śmieszny
- wciągający

Następnie tworzymy wektor dla każdego tekstu obrazujący wystąpienia słów ze słownika w tekstach.

TEKSTY \ SŁOWNIK	film	jest	i	bardzo	długi	pouczający	śmieszny	wciągający	przerażający
<i>Film jest bardzo przerażający i długi.</i>	1	1	1	1	1	0	0	0	1
<i>Film jest przerażający i jest pouczający.</i>	1	2	1	0	0	1	0	0	1
<i>Film jest śmieszny i wciągający.</i>	1	1	1	0	0	0	1	1	0

Unormowanie

TEKSTY \ SŁOWNIK	film	jest	i	bardzo	długi	pouczający	śmieszny	wciągający	przeróżający
<i>Film jest bardzo przerażający i długi.</i>	1/6	1/6	1/6	1/6	1/6	0	0	0	1/6
<i>Film jest przerażający i jest pouczający.</i>	1/6	1/6	1/6	0	0	1/6	0	0	1/6
<i>Film jest śmieszny i wciągający.</i>	1/5	1/5	1/5	0	0	0	1/5	1/5	0

Deskryptory cech

Deskryptor cech (ang. feature descriptor) - algorytm, który produkuje liczbowy opis wcześniej znalezionych cech na obrazku. Zazwyczaj jest to wektor lub zbiór wektorów.

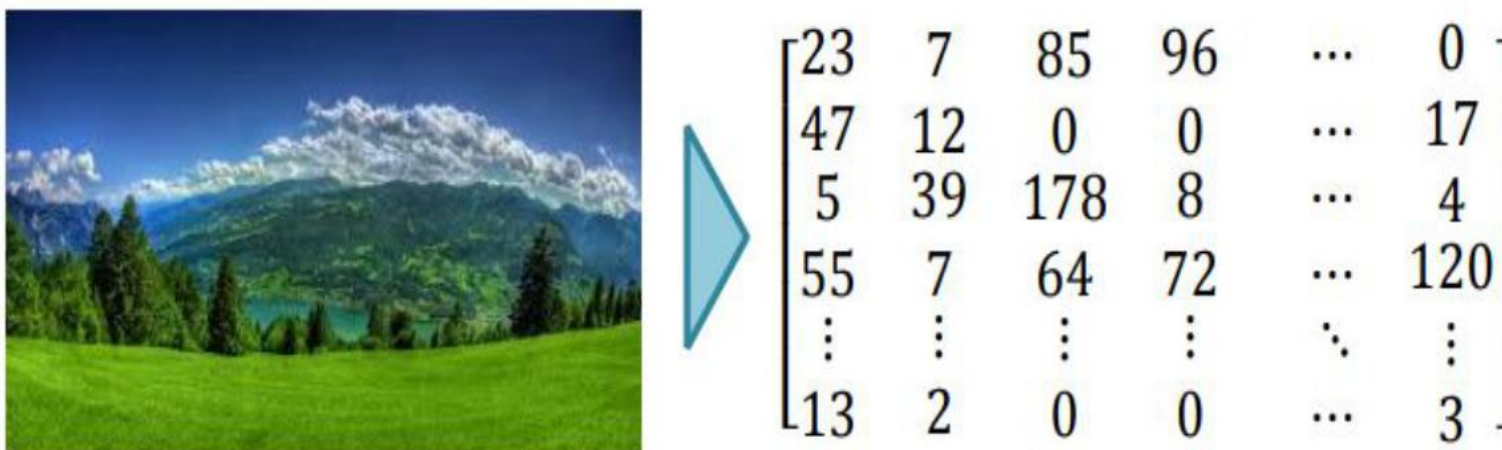


Figure 1: Przekształcenie obrazu do postaci numerycznej. Wykorzystując cechy lokalne otrzymuje się macierz o liczbie kolumn równej długości deskryptora i liczbie wierszy równej liczbie znalezionych punktów kluczowych

Algorytm k -średnich

Algorytm ma następujące kroki:

- 1 określamy liczbę klastrów K
- 2 ustalamy początkowe centra klastrów m_1, m_2, \dots, m_K , zazwyczaj są one losowo wybranymi obserwacjami z danych uczących

- 3 przyporządkowujemy elementy x_1, \dots, x_N do najbliższego centroidu:

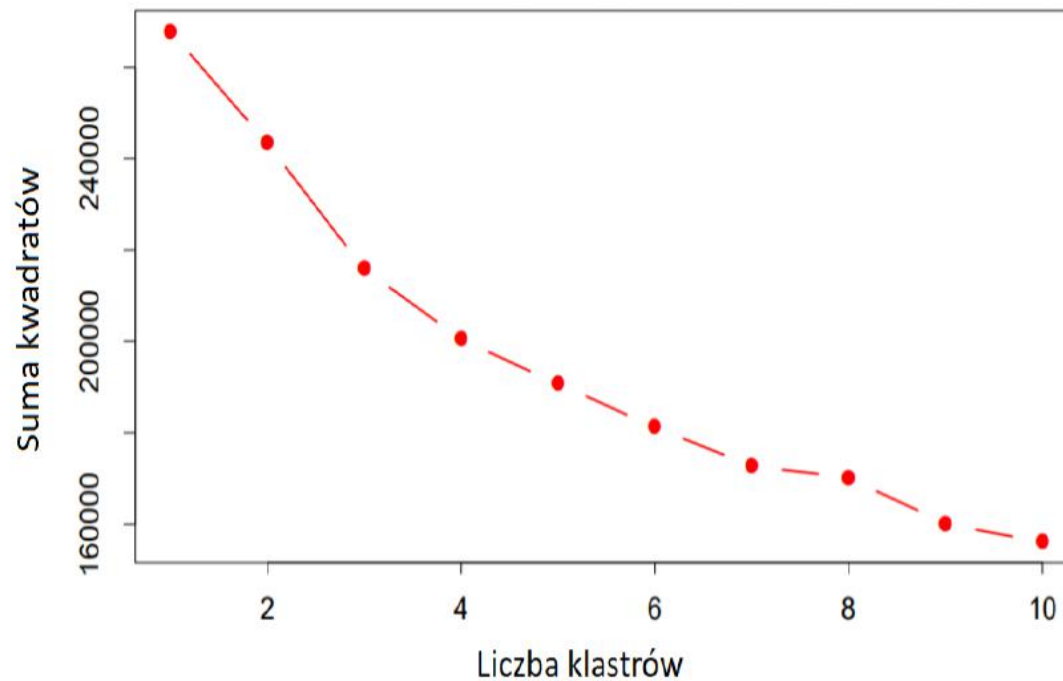
$$C(i) = \arg \min_{1 \leq k \leq K} \|x_i - m_k\|$$

- 4 wyznaczamy nowe środki skupień m_1, m_2, \dots, m_K będące średnią arytmetyczną po współrzędnych elementów skupienia
- 5 kroki 3 i 4 są powtarzane do osiągnięcia limitu iteracji lub do momentu, w którym przypisania elementów do klastrów się nie zmieniają.

K średnich, metoda „łokcia”

Dla każdej liczby klas K obliczamy

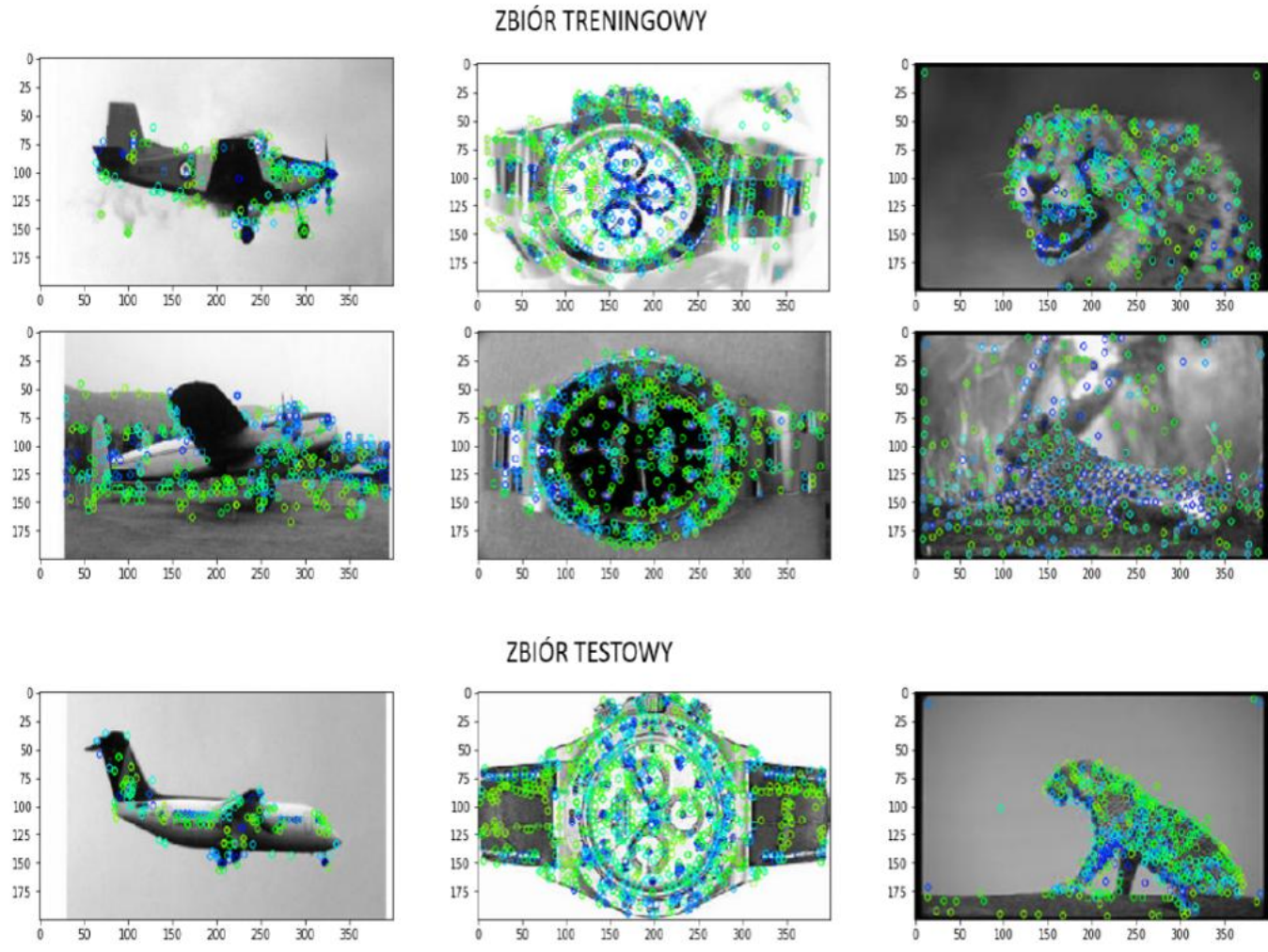
$$\sum_{k=1}^K \sum_{x_i \in C_k} \|x_i - m_k\|^2$$



Bag of features

Tę samą metodologię co w worku słów aplikuje się do obrazów cyfrowych poprzez wprowadzenie pojęcia „wizualnego słowa” będącego analogią „słowa” w dokumencie tekstowym. Wizualne słowa są deskryptorami cech lokalnych. Worek cech jest zatem histogramem częstości występowania wizualnych słów w obrazie.

Problem



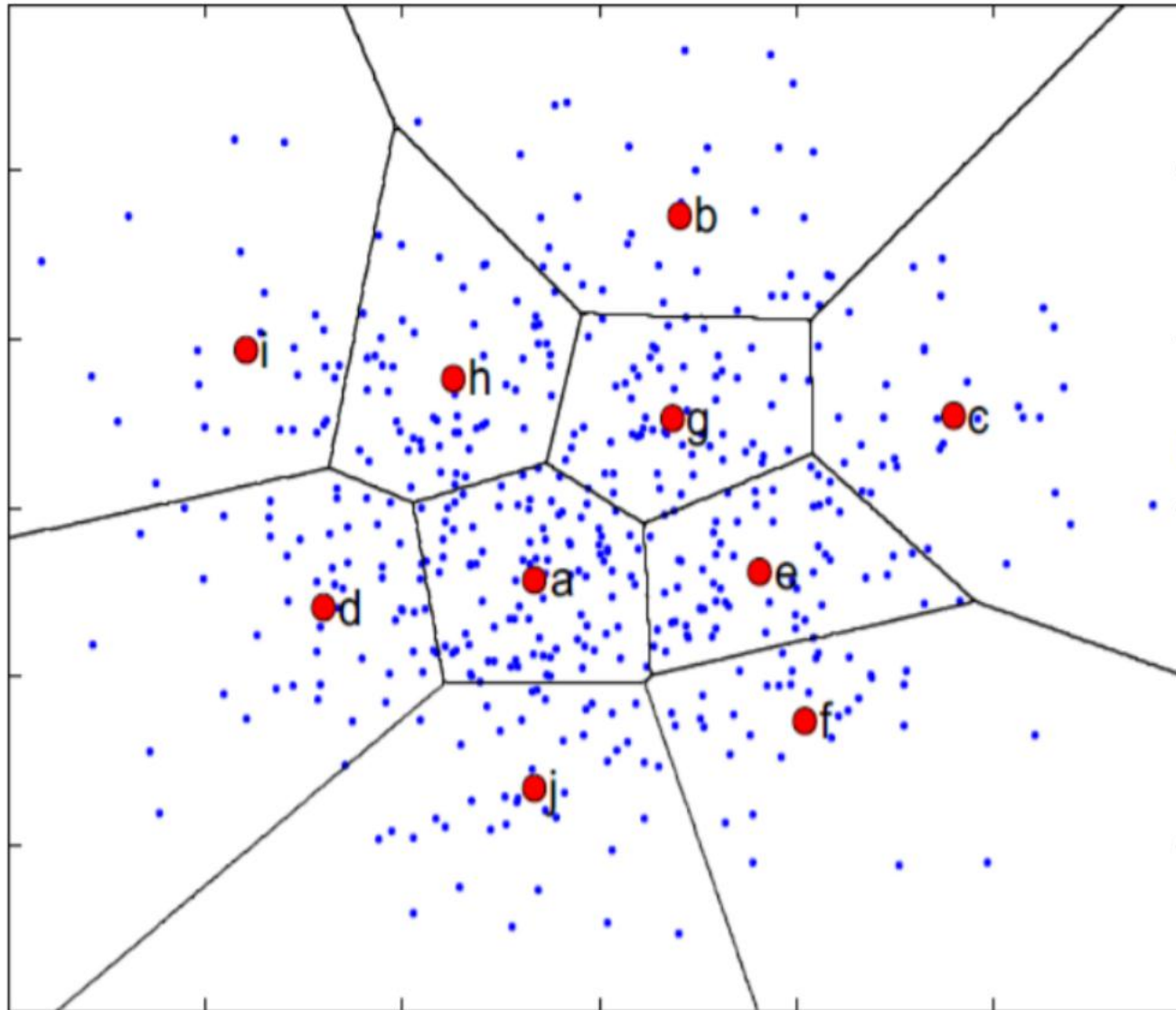
Konstrukcja słownika

Rozważmy zbiór obrazków treningowych $\{O_1, O_2, \dots, O_N\}$ pochodzących z S kategorii. Na obrazku O_i znaleziono algorytmem SIFT wektory $\{x_1^i, \dots, x_{n_i}^i\}$, gdzie $x^i \in R^{128}$. Do skonstruowania słownika używamy jako danych wszystkich punktów ze wszystkich obrazków tzn. $\cup_{j=1}^N \cup_{i=1}^{n_j} x_j^i$. Dla takich danych aplikujemy algorytm k -średnich. Każdy wektor x_j^i przyporządkujemy do najbliższego centroidu tzn.

$$C(i, j) = \arg \min_{1 \leq k \leq K} \|x_j^i - m_k\|,$$

gdzie $\{m_1, m_2, \dots, m_K\}$ to centra klastrów. Liczba klas może zostać wybrana np. „regułą łokcia”. Po wybraniu odpowiedniej liczby klas K , $\{m_1, m_2, \dots, m_K\}$ tworzą słownik. Centroidy to reprezentatywne deskryptory zwane wizualnymi słowami.

Konstrukcja słownika



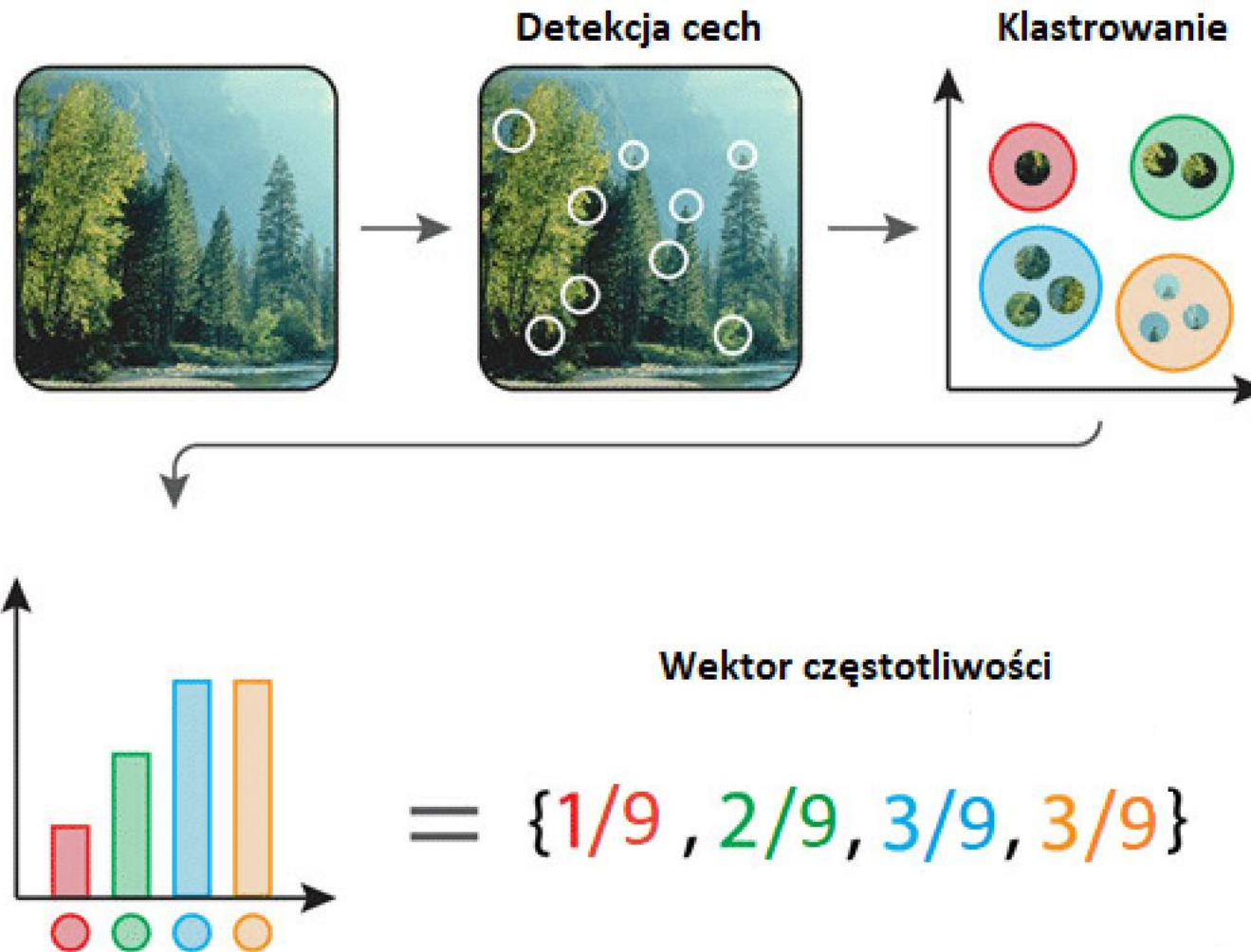
Wyznaczanie histogramu

$$f^i = (f_1^i, f_2^i, \dots, f_K^i) = \left(\frac{\sum_{w=1}^{n_i} \mathbf{1}_{\{C(i,w)=1\}}}{n_i}, \frac{\sum_{w=1}^{n_i} \mathbf{1}_{\{C(i,w)=2\}}}{n_i}, \dots, \frac{\sum_{w=1}^{n_i} \mathbf{1}_{\{C(i,w)=K\}}}{n_i} \right)$$

Takie wektory konstruujemy dla wszystkich obrazków. Finalnie wynikiem tego kroku jest zbiór wektorów

$$\{f^1, f^2, \dots, f^N\}.$$

Wyznaczanie histogramu



Klasyfikacja

Dla zdjęć ze zbioru testowego $\{O_1^T, O_2^T, \dots, O_P^T\}$, dla każdego obrazka O_i^T znajdujemy algorytmem SIFT wektory $\{v_1^i, \dots, v_{p_i}^i\}$, gdzie $v^i \in R^{128}$. Korzystając z utworzonego uprzednio słownika obliczamy

$$C^T(i, j) = \arg \min_{1 \leq k \leq K} \|v_j^i - m_k\|.$$

Następnie konstruujemy wektory częstotliwości

$$g^i = (g_1^i, g_2^i, \dots, g_K^i) = \left(\frac{\sum_{w=1}^{p_i} \mathbf{1}_{\{C^T(i,w)=1\}}}{p_i}, \frac{\sum_{w=1}^{p_i} \mathbf{1}_{\{C^T(i,w)=2\}}}{p_i}, \dots, \frac{\sum_{w=1}^{p_i} \mathbf{1}_{\{C^T(i,w)=K\}}}{p_i} \right)$$

Tak uzyskujemy zbiór wektorów

$$\{g^1, g^2, \dots, g^P\}.$$

Modyfikacja worka cech - redukcja punktów kluczowych

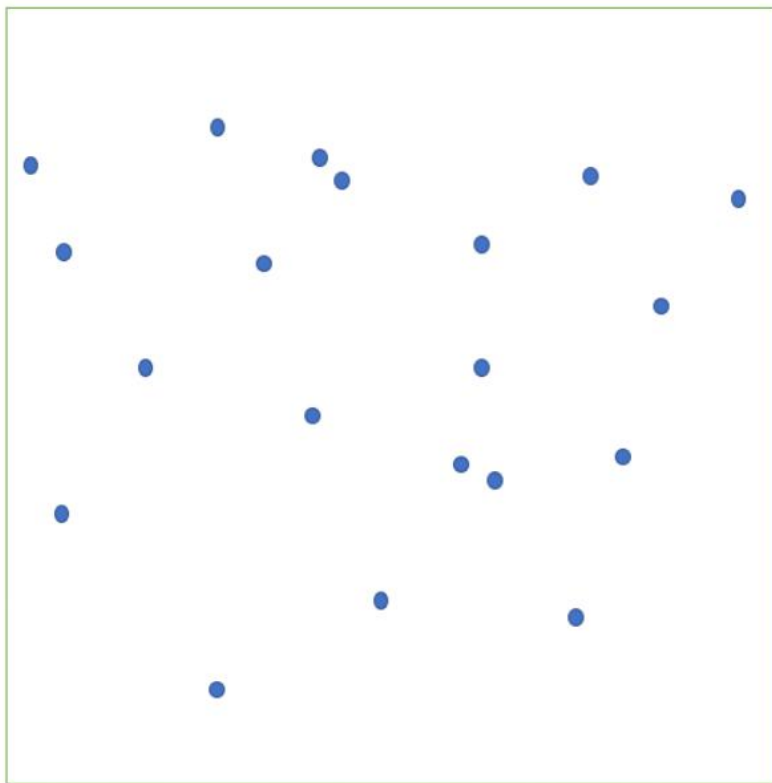
Problem

W przypadku analiz dużych zbiorów obrazów liczba punktów znalezionych przez algorytm SIFT może być trudna do przetworzenia przez komputer i zastosowania w klasyfikacji.

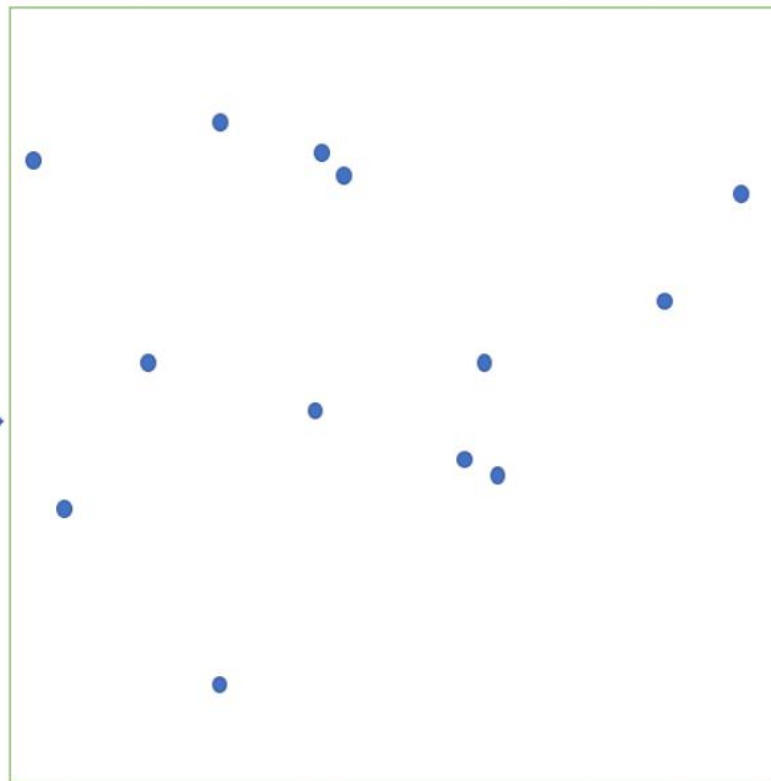
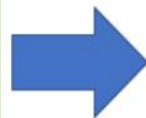
Rozwiązanie

Redukcja punktów znalezionych przez algorytm SIFT, tak aby przyspieszyć szybkość obliczeń i ograniczyć liczbę zapamiętywanych danych bez pogorszenia wyników klasyfikacji.

METODA M1 – Redukcja punktów kluczowych losowo

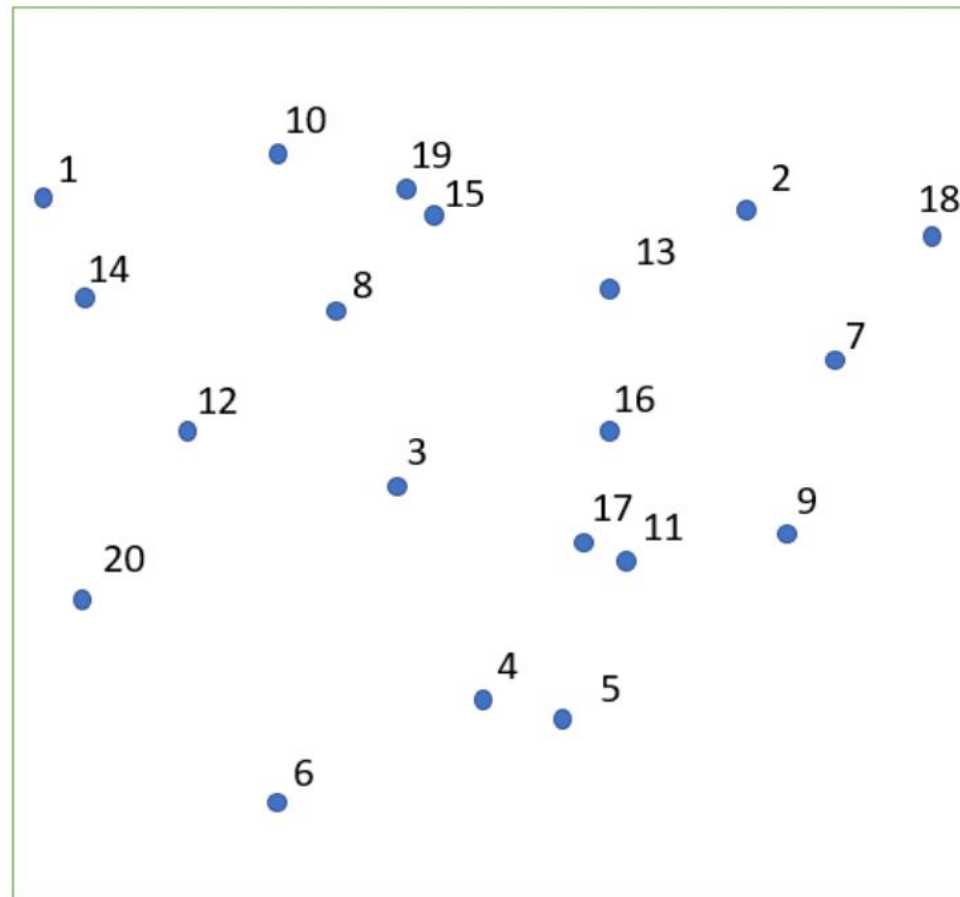


OBRAZ PRZED REDUKCJĄ



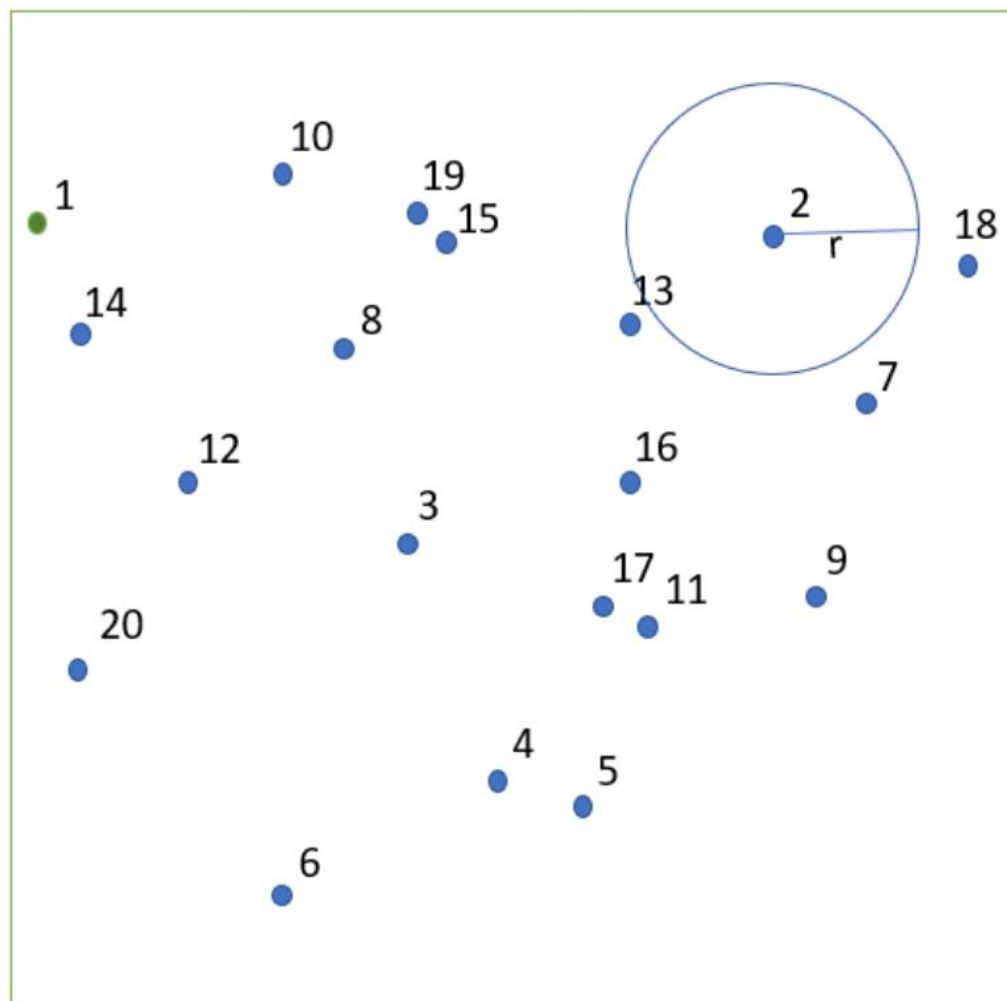
OBRAZ PO REDUKCJĄ

METODA M2 – Redukcja punktów kluczowych według odległości i losowej permutacji

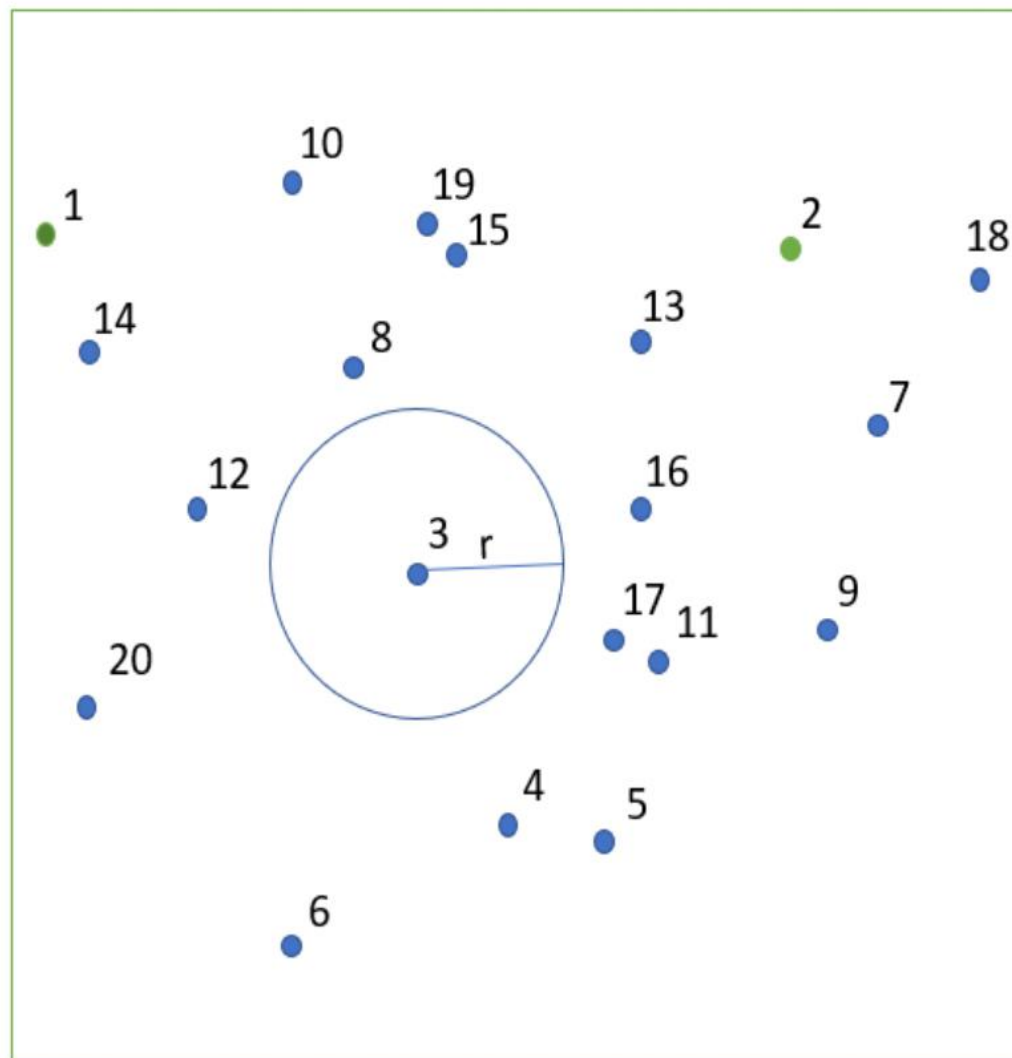


OBRAZ PRZED REDUKCJĄ

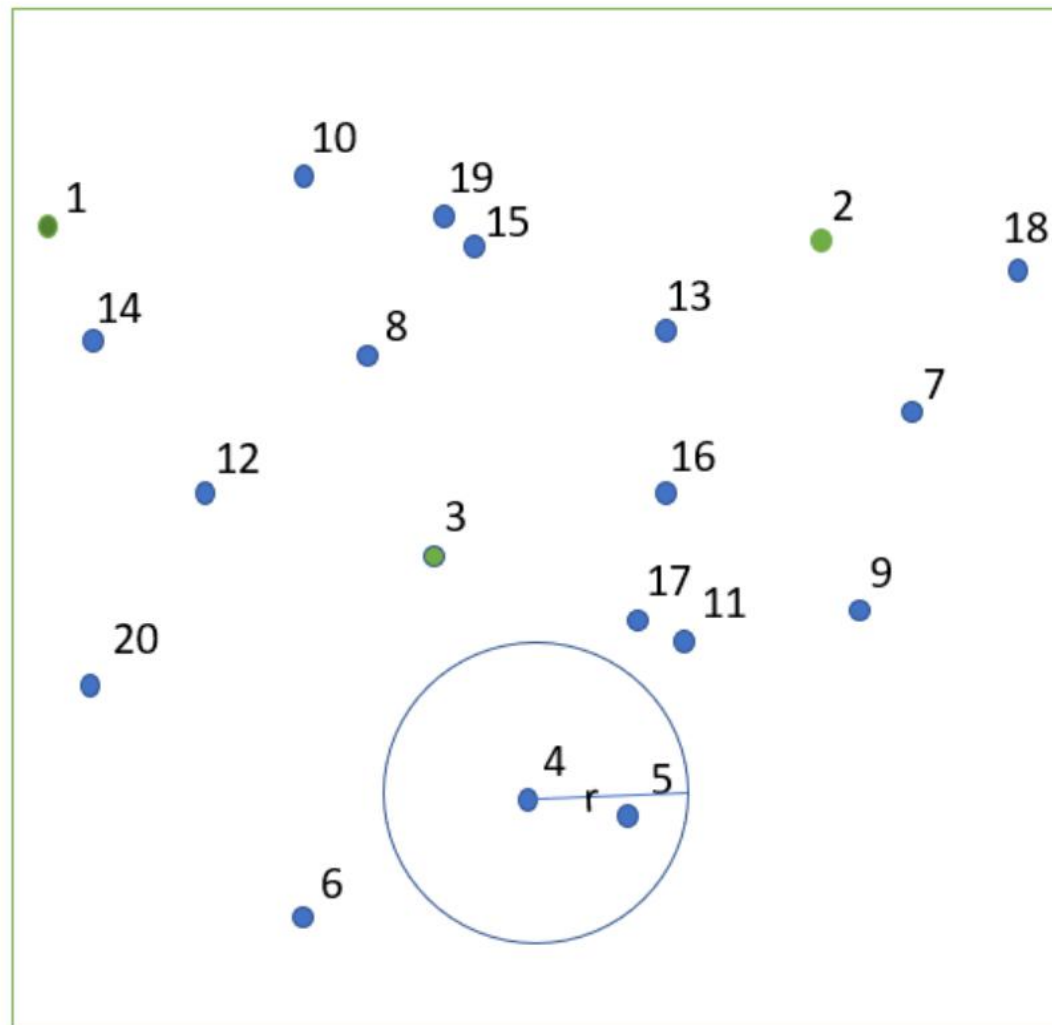
METODA M2 – Redukcja punktów kluczowych według odległości i losowej permutacji



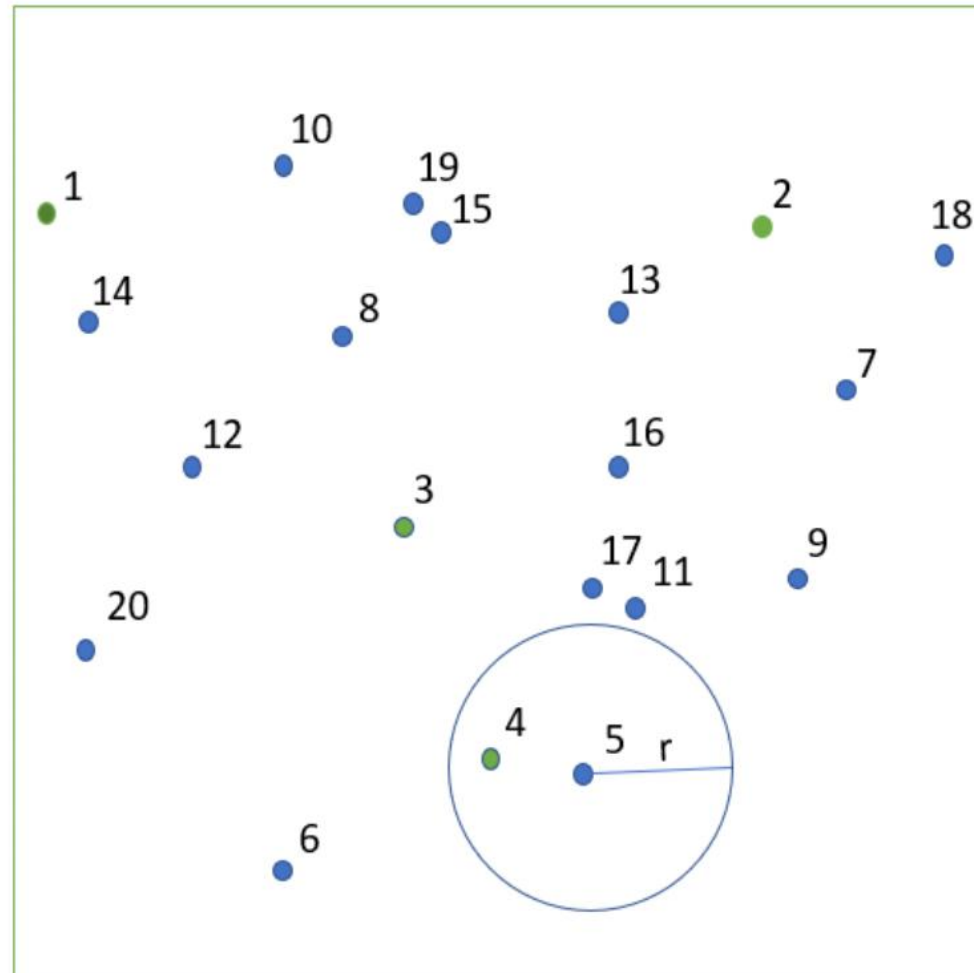
METODA M2 – Redukcja punktów kluczowych według odległości i losowej permutacji



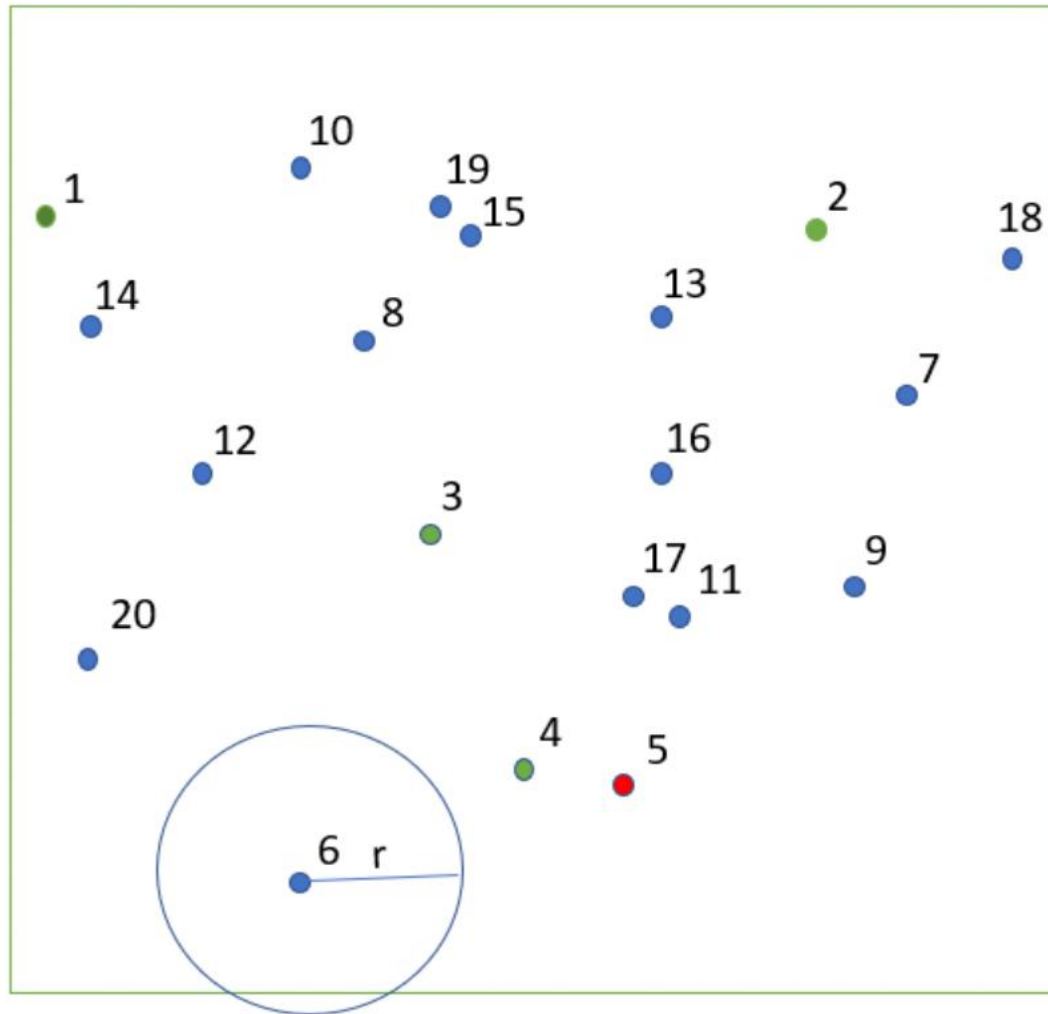
METODA M2 – Redukcja punktów kluczowych według odległości i losowej permutacji



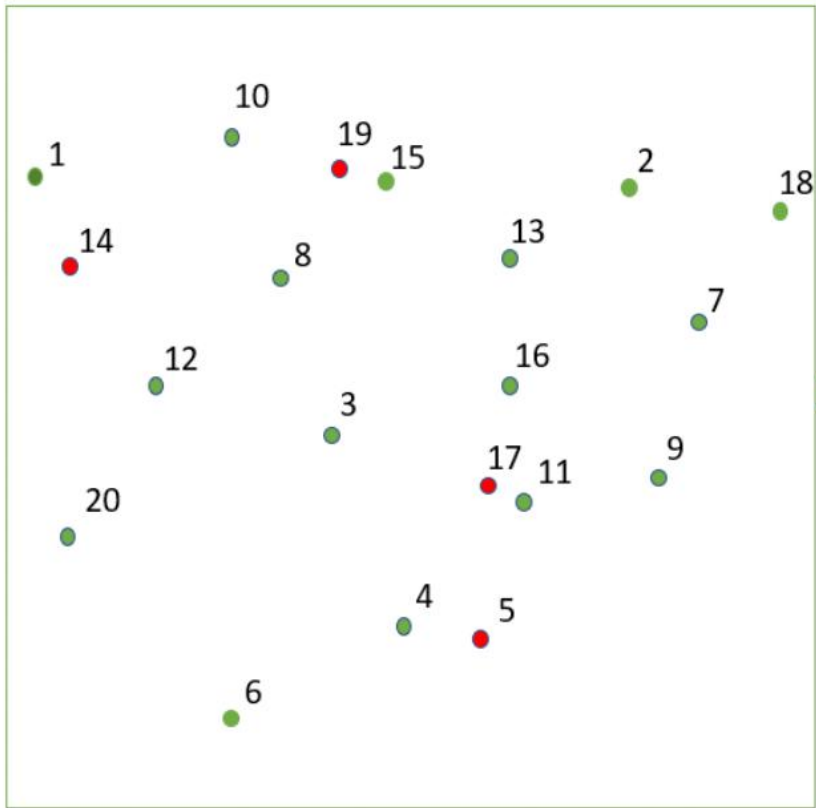
METODA M2 – Redukcja punktów kluczowych według odległości i losowej permutacji



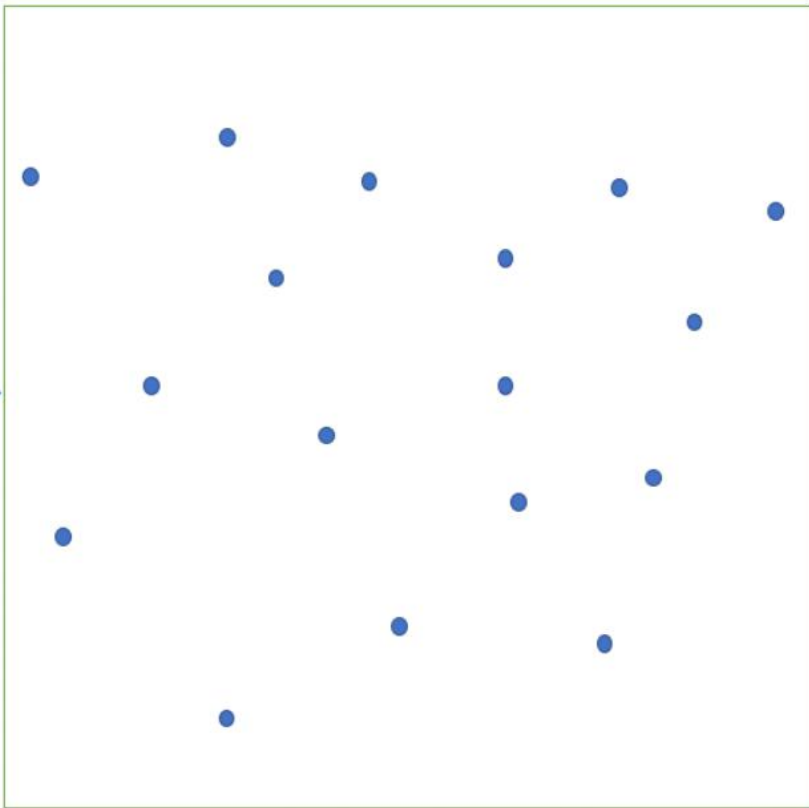
METODA M2 – Redukcja punktów kluczowych według odległości i losowej permutacji



METODA M2 – Redukcja punktów kluczowych według odległości i losowej permutacji



OBRAZ PRZED REDUKCJĄ



OBRAZ PO REDUKCJI

METODA M3 – Redukcja punktów kluczowych według odległości i mocy

Metoda M3, czyli modyfikacja M2

Nie ustawiamy punktów według losowej permutacji, tylko według ich mocy.

Wyniki pracy

Table 1: Wyniki klasyfikacji dla metody M1

metoda	nie usuwamy punktów	usuwamy 30% punktów według M1	usuwamy 20% punktów według M1
SVM	79%	78%	78%
<i>k</i> -NN	61%	58%	56%
las losowy	75%	74%	74%
Naiwny Bayes	78%	78%	78%

Wyniki pracy

Table 2: Wyniki klasyfikacji dla metody M2

metoda	nie usuwamy punktów	usuwamy 30% punktów według M2	usuwamy 20% punktów według M2
SVM	79%	78%	79%
k -NN	61%	58%	59%
las losowy	75%	76%	79%
Naiwny Bayes	78%	78%	78%

Wyniki pracy

Table 3: Wyniki klasyfikacji dla metody M3.

metoda	nie usuwamy punktów	usuwamy 30% punktów według M3	usuwamy 20% punktów według M3
SVM	79%	80%	79%
k-NN	61%	61%	63%
las losowy	75%	75%	76%
Naiwny Bayes	78%	78%	78%