

Магістерська кваліфікаційна робота на тему:

# ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ЛІНГВІСТИЧНОЇ ІНДЕКСАЦІЇ КАРТИН НА ОСНОВІ НЕЙРОННОЇ МЕРЕЖІ

Виконав магістрант гр. 2КН-16м Машталір І.В.  
Керівник: к.т.н., доц. Колесницький О.К.

# МЕТА ТА ЗАВДАННЯ ДОСЛІДЖЕННЯ

Метою дослідження магістерської кваліфікаційної роботи є підвищення достовірності лінгвістичної індексації картин за рахунок використання інформаційних технологій штучних нейронних мереж.

**Для досягнення поставленої мети необхідно розв'язати такі завдання:**

- провести аналіз проблеми розв'язання задачі лінгвістичної індексації картин;
- розглянути існуючі методи вирішення задачі лінгвістичної індексації картин та обрати й обґрунтувати вибір методу, який задовольняє мету даної магістерської кваліфікаційної роботи;
- розробити математичну модель лінгвістичної індексації картин;
- розробити структуру інформаційної системи лінгвістичної індексації картин;
- сформулювати стадії інформаційної технології, розробити структуру та алгоритм роботи програмного засобу;
- виконати програмну реалізацію запропонованої інформаційної технології;
- провести тестування програмного продукту та виконати аналіз отриманих результатів.

# ОБ'ЄКТ, ПРЕДМЕТ ТА МЕТОДИ ДОСЛІДЖЕННЯ

Об'єкт дослідження – процес лінгвістичної індексації (класифікації по змісту) картин з використанням інформаційних технологій.

Предмет дослідження – інформаційна технологія та програмні засоби лінгвістичної індексації картин з використанням нейронної мережі та достовірність індексації при їх використанні.

## Методи дослідження

У роботі використані наступні методи наукових досліджень:

- системного аналізу для аналізу структури інформаційної системи,
- теорії нейронних мереж для реалізації інформаційної технології лінгвістичної індексації картин
- теорія нейронних мереж для реалізації інформаційної технології,
- методи математичної статистики для розробки процесу лінгвістичної індексації картин та обрахунків результатів експериментів із програмним засобом,
- об'єктно-орієнтованого програмування для програмної реалізації.

# НАУКОВА НОВИЗНА ОДЕРЖАНИХ РЕЗУЛЬТАТІВ

полягає в наступному:

- набула подальшого розвитку інформаційна технологія лінгвістичної індексації картин, яка відрізняється використанням для виділення ознак методу головних компонент, а для класифікації ознак - штучної нейронної мережі зворотного поширення, що дозволило підвищити достовірність лінгвістичної індексації картин;
- удосконалено метод головних компонент для виділення ознак картини, який відрізняється використанням колірному простору зображення YCbCr замість RGB, що дозволило зменшити потік відеоданих, а значить збільшити швидкість індексації картин.

## ПРАКТИЧНЕ ЗНАЧЕННЯ ОДЕРЖАНИХ РЕЗУЛЬТАТІВ

1. Розроблено алгоритм лінгвістичної індексації картин на основі нейронної мережі.
  2. Розроблено алгоритм виділення ознак зображення на основі методу головних компонент.
  3. Розроблено програмний засіб для лінгвістичної індексації картин на основі нейронної мережі зворотного поширення.
- Розроблені алгоритми можуть бути впроваджені в початковий процес як лекція на тему «Нейромережевий метод лінгвістичної індексації картин» дисципліни «Нейромережеві методи обчислювального інтелекту».

# Постановка задачі лінгвістичної індексації картин

потрібно графічному зображенню поставити у відповідність текстову назву класу, до якого воно відноситься, або іншими словами – лінгвістичний індекс цього класу. Таким чином, означену задачу класифікації картин за змістом можна назвати як задача лінгвістичної індексації картин.

Програма, яка розробляється, повинна виконувати такі функції:

вибір файлу картини у заданому розташуванні та завантаження його в програму для обробки;

відображення обраного файлу у вікні програми;

визначення класу, до якого належить завантажена картина;

кількість класів повинна бути не менше 5;





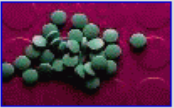
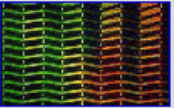



навчання програми класифікувати картини на основі підготовлених прикладів;

## Аналіз методів лінгвістичної індексації картин



Було обрано метод головних компонент та класифікатор на базі штучних нейронних мереж

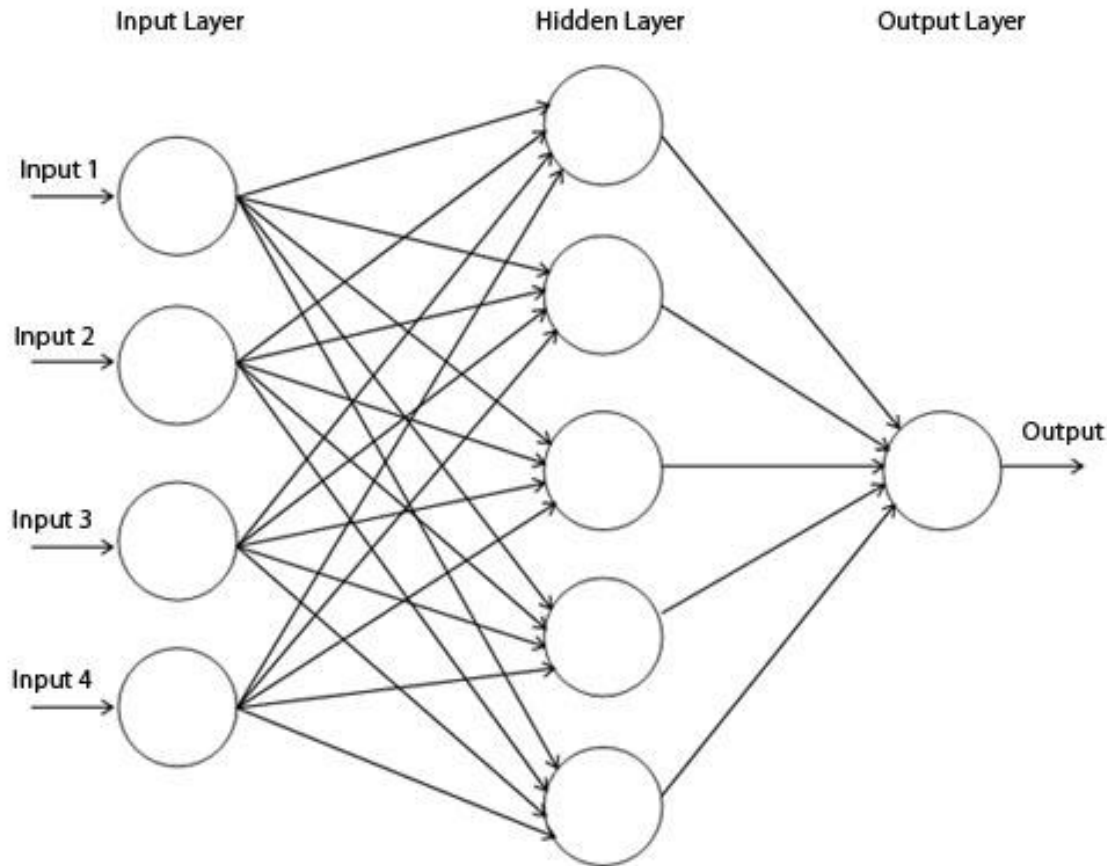
# Вибір і обґрунтування аналогу

 <p><b>Computer Predictions –</b> :ocean: :wave: :fun: :wind: :ski: :ice: :beach: :snow: :wild_life: :sport: :mountain: :sky: :water: :landscape: :animal: :people:</p> <p>Manual Category Annotation – :painting: :man-made: :landscape:</p>	 <p><b>Computer Predictions –</b> :historical_building: :London: :Ottawa: :Pyramid: :Silkroad: :church: :Asia: :sky: :water: :people: :landscape:</p> <p>Manual Category Annotation – :New_York: :city: :sky_line: :building: :modern:</p>	 <p><b>Computer Predictions –</b> :man-made: :ballet: :doll: :monument: :child: :indoor: :ski: :old: :plane: :ice: :snow: :sport: :sky: :cloth: :people: :ski: :man-made: :snow: :ice: :sport: :cloth:</p> <p>Manual Category Annotation –</p>
 <p><b>Computer Predictions –</b> :car: :London: :pomp_pag: :parade: :bus: :England: :food: :Asia: :boat: :man-made: :sky: :cloth: :people: :building: :historical_building: :England: :landscape: :mountain: :lake: :tree:</p> <p>Manual Category Annotation –</p>	 <p><b>Computer Predictions –</b> :fractal: :jewelry: :pill: :color: :man-made: :indoor: :pill: :man-made: :texture:</p> <p>Manual Category Annotation –</p>	 <p><b>Computer Predictions –</b> :dusk: :dawn: :game: :Finland: :light: :night: :sky: :city: :sport: :mountain: :Europe: :building: :man-made: :historical_building: :people: :landscape: :color: :texture: :man-made:</p> <p>Manual Category Annotation –</p>
 <p><b>Computer Predictions –</b> :England: :everglade: :house: :lake: :race: :plant: :grass: :horse: :car: :bird: :flower: :animal: :man-made: :people: :landscape: :London: :historical_building: :sky: :water: :people:</p> <p>Manual Category Annotation –</p>	 <p><b>Computer Predictions –</b> :sun: :sky: :cloud: :dawn: :dusk: :orange: :plant: :landscape: :rural: :flower: :plant: :England:</p> <p>Manual Category Annotation –</p>	 <p><b>Computer Predictions –</b> :indoor: :old: :ballet: :dish: :antique: :child: :thing: :man-made: :people: :old: :toy: :man-made: :indoor:</p> <p>Manual Category Annotation –</p>

Аналог - програма  
ALIPR (Automatic Linguistic Indexing  
of Pictures - Real Time)

# ОБГРУНТУВАННЯ ВИБОРУ НЕЙРОННОЇ МЕРЕЖІ

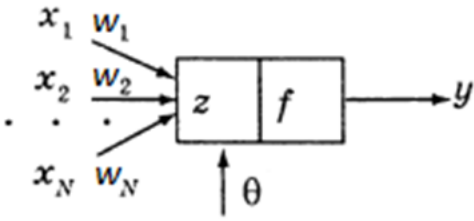
Була обрана нейронна мережа багатозаровий перцептрон, що навчається за допомогою методу зворотного поширення помилки,



Структура багатозарового персептрона

# Математична модель багатозарового персептрона

## Метод зворотного поширення помилки



Структура формального нейрона

$$z = \sum_{i=0}^N x_i w_i, \quad (1)$$

$$f(z) = \begin{cases} 1, & z \geq \theta \\ 0, & z < \theta \end{cases}$$

$$f_i^{(j)} = f\left(\sum_{t=1}^{N^{(j-1)}} f_t^{(j-1)} w_{ti}^{(j)} - \theta_i^{(j)}\right) \quad (2)$$

Використання квадратичного критерію якості навчання:

дозволяє отримати градієнтний алгоритм корекції ваг:

Похибка поширюється від виходів до входів, тому спочатку обчислюються корекції ваг для вихідного шару.:

Отримані складові виводимо наступним чином:

Виділимо складову k-го шару окремо:

Підставивши (8) – (11) в (7) отримуємо для вихідного шару:

$$\varepsilon_i^2 = (y_i - y_i^*)^2 \quad (3)$$

$$\Delta w_{ij}^{(k)} = -\gamma_{ij} \cdot \frac{\partial \varepsilon_j^2}{\partial w_{ij}^{(k)}} \quad (4)$$

де  $\gamma_{ij}$  - коефіцієнт швидкості навчання.

$$\frac{\partial \varepsilon_j^2}{\partial w_{ij}^{(k)}} = \frac{\partial \varepsilon_j^2}{\partial f_j^{(k)}} \cdot \frac{\partial f_j^{(k)}}{\partial z_j^{(k)}} \cdot \frac{\partial z_j^{(k)}}{\partial w_{ij}^{(k)}} \quad (5)$$

$$\Delta w_{ij}^{(k)} = 2\varepsilon_i \gamma_{ij} \cdot \alpha f_j^{(k)} (1 - f_j^{(k)}) \cdot f_i^{(k-1)} \quad (6)$$

$$\delta_j^{(k)} = 2\varepsilon_j \alpha f_j^{(k)} (1 - f_j^{(k)}) \quad (7)$$

$$\Delta w_{ij}^{(k)} = \gamma_{ij} \cdot \delta_j^{(k)} \cdot f_i^{(k-1)} \quad (8)$$

$$\frac{\partial \varepsilon_j^2}{\partial f_j^{(k)}} = -2(y_j - y_j^*) = -2\varepsilon_j \quad (9)$$

$$\frac{\partial f_j^{(k)}}{\partial z_j^{(k)}} = \alpha f_j^{(k)} (1 - f_j^{(k)}) \quad (10)$$

$$\frac{\partial z_j^{(k)}}{\partial w_{ij}^{(k)}} = f_i^{(k-1)} \quad (11)$$

$$\Delta w_{ij}^{(k)} = 2\varepsilon_i \gamma_{ij} \cdot \alpha f_j^{(k)} (1 - f_j^{(k)}) \cdot f_i^{(k-1)} \quad (12)$$



# Математична модель методу головних компонент

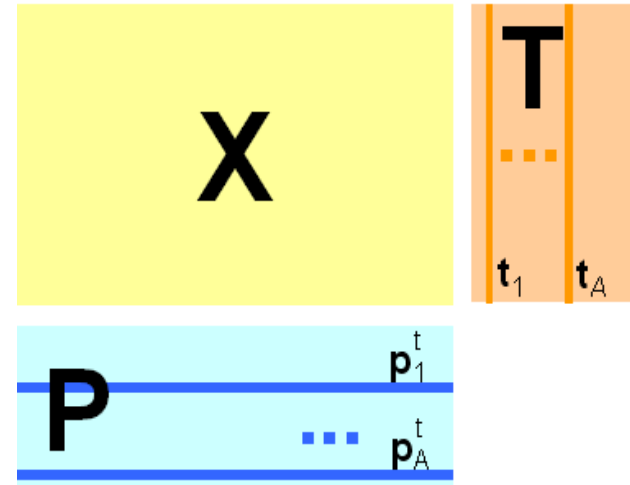
Суть методу головних компонент - це істотне зниження розмірності даних. Вихідна матриця  $X$  замінюється двома новими матрицями  $T$  і  $P$ , розмірність яких,  $A$ , менше, ніж число змінних (стовпців)  $J$  у вихідній матриці  $X$

$$X = \begin{matrix} \begin{matrix} x_{11} & x_{12} & \dots & \dots & x_{1j} & & x_{1J} \\ x_{21} & x_{22} & \dots & \dots & x_{2j} & & x_{2J} \\ \dots & \dots & \dots & \dots & \dots & & \dots \\ x_{i1} & x_{i2} & \dots & \dots & x_{ij} & & x_{iJ} \\ \dots & \dots & \dots & \dots & \dots & & \dots \\ x_{J1} & x_{J2} & \dots & \dots & x_{Jj} & & x_{JJ} \end{matrix} \end{matrix}$$

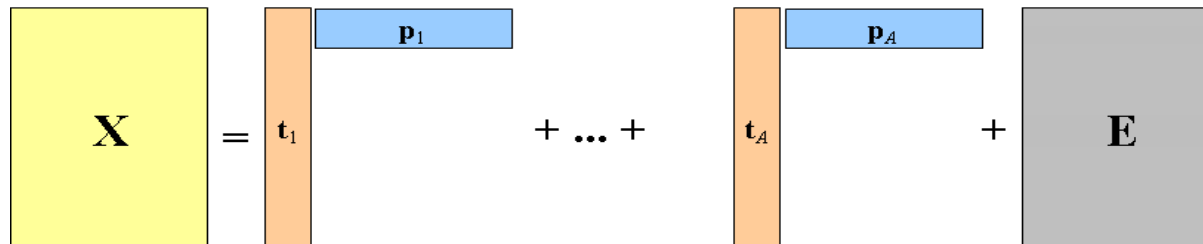
Вихідна матриця даних розмірністю  $I$  рядків і  $J$  стовпців

$$X = TP^t + E = \sum_{a=1}^A t_a p_a^t + E$$

За допомогою цих нових змінних матриця  $X$  розкладається в добуток двох матриць  $T$  і  $P$



Декомпозиція матриці  $X$



Розкладання по головних компонентах

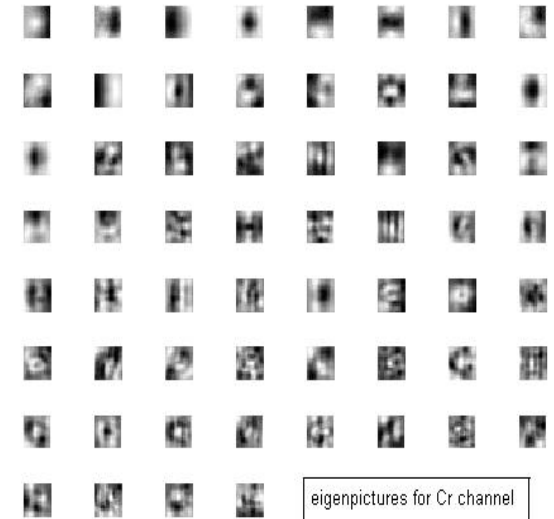
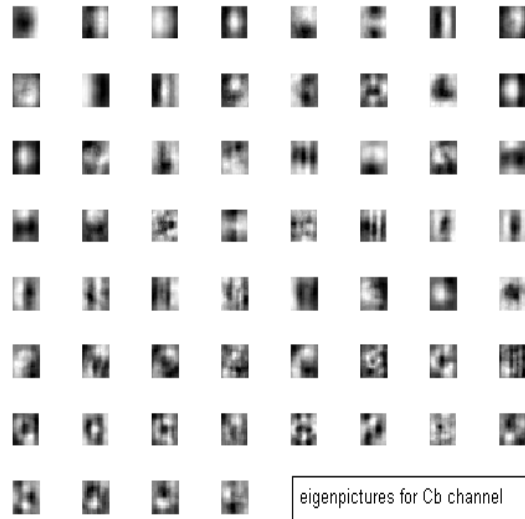
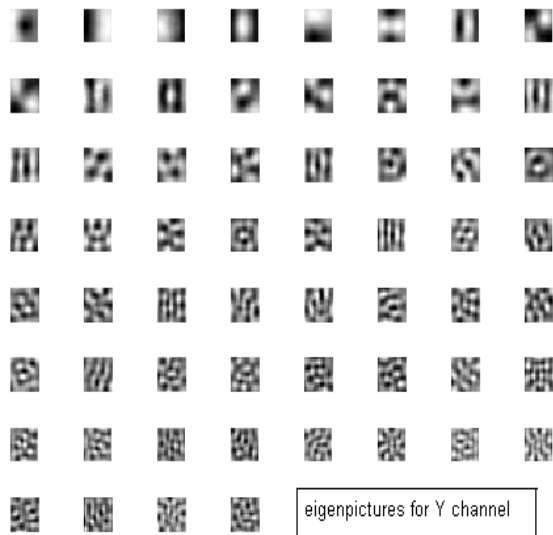
# РОЗРОБКА СТРУКТУРНОЇ СХЕМИ СИСТЕМИ ЛІНГВІСТИЧНОЇ ІНДЕКСАЦІЇ КАРТИН



# АЛГОРИТМ ВИДІЛЕННЯ ОЗНАК ЗОБРАЖЕННЯ

У цій роботі було обрано найпростіші ознаки зображення, які не схожі на якісь раціональні ознаки взагалі. Як ознаки було обране саме зображення, розмір якого було знижено до 16x16 пікселів і перетворене в колірний формат YCbCr.

YCbCr - альтернативний колірний простір, який отримується з RGB. Y - компонента яскравості, Cb і Cr є синьою і червоною кольорорізницевиими компонентами.



Перші 60 власних зображень  
для каналу Y колірного  
простору YCbCr

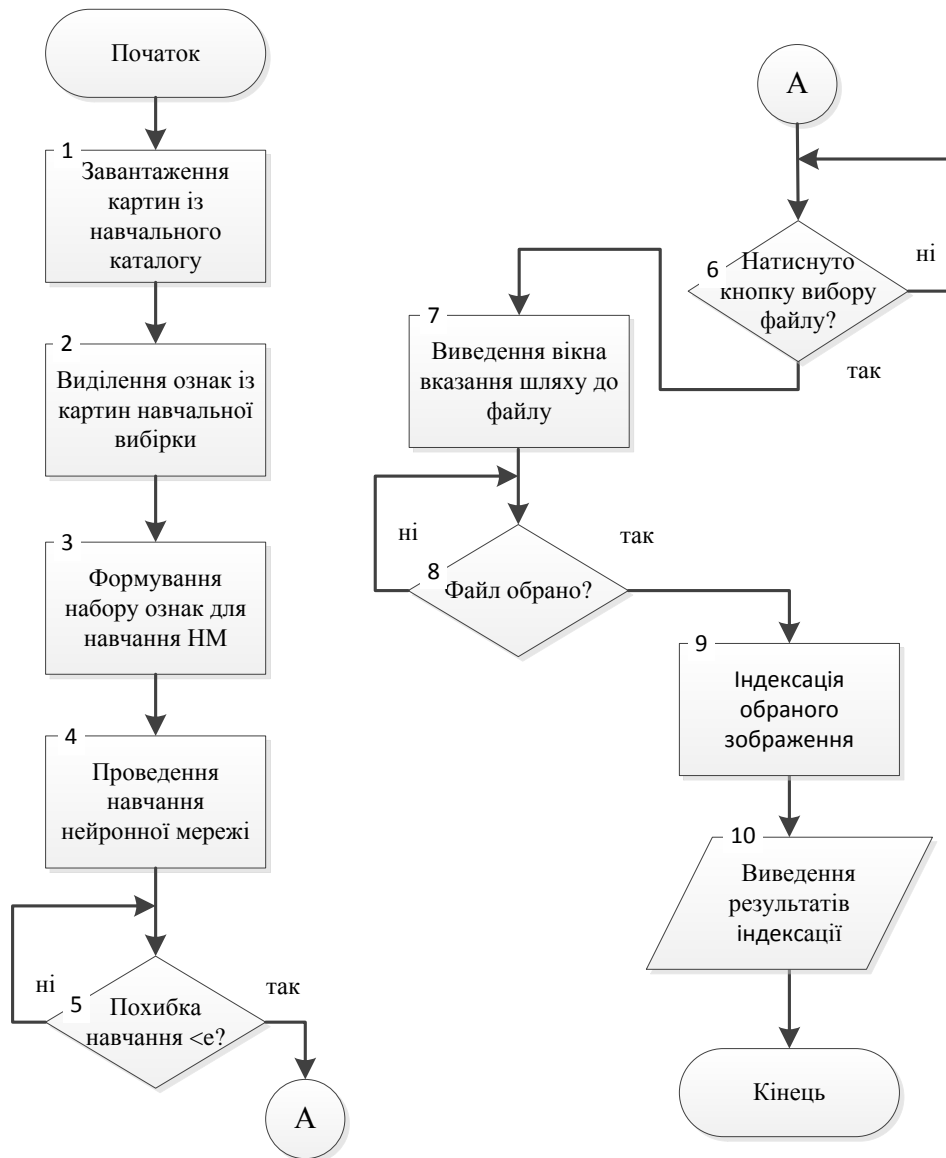
Перші 60 власних зображень  
для каналу Cb колірного  
простору YCbCr

Перші 60 власних зображень  
для каналу Cr колірного  
простору YCbCr

# Структура інформаційної технології лінгвістичної індексації картин на основі нейронної мережі



# АЛГОРИТМ ЛІНГВІСТИЧНОЇ ІНДЕКСАЦІЇ КАРТИН

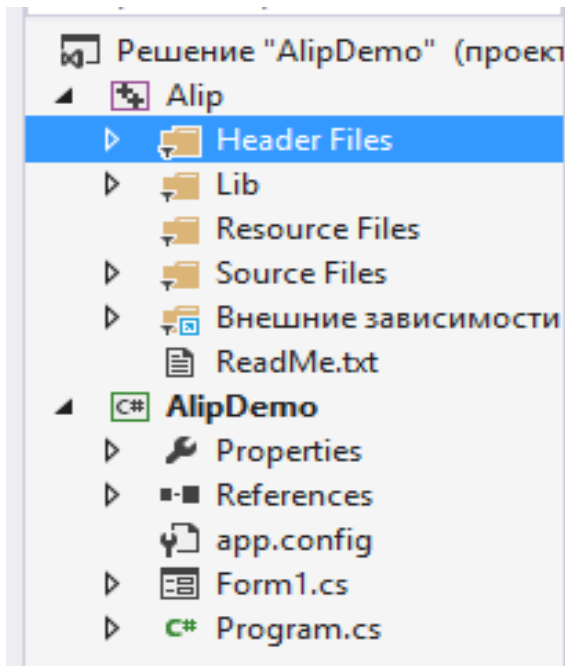


Для визначеності, при розробці програми передбачено 5 простих категорій (класів) природніх зображень:

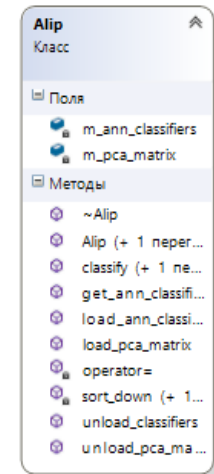
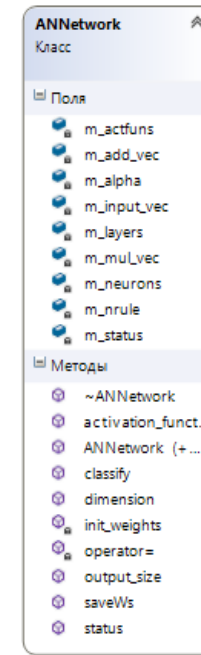
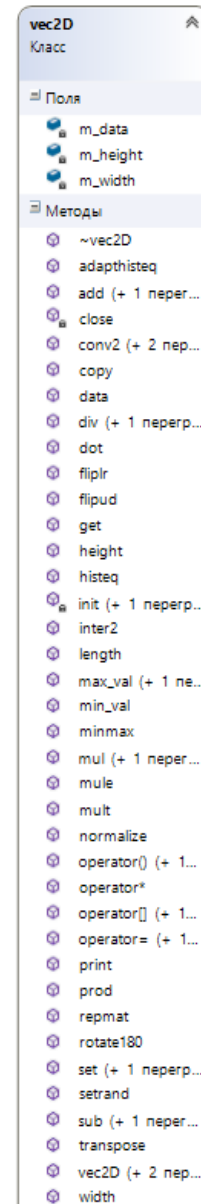
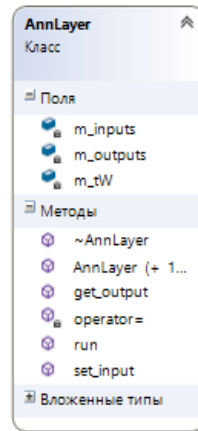
- Зображення, які можуть містити тварин.
- Зображення, які можуть містити квіти.
- Зображення, які можуть містити ландшафти.
- Зображення, які можуть містити сходи та заходи сонця.
- Інші зображення, які не містять вищевказані категорії або просто невідомий тип зображення.

Для програмної реалізації інформаційної технології лінгвістичної індексації картин було обрано мову програмування C#

# Розробка структури програмного додатку



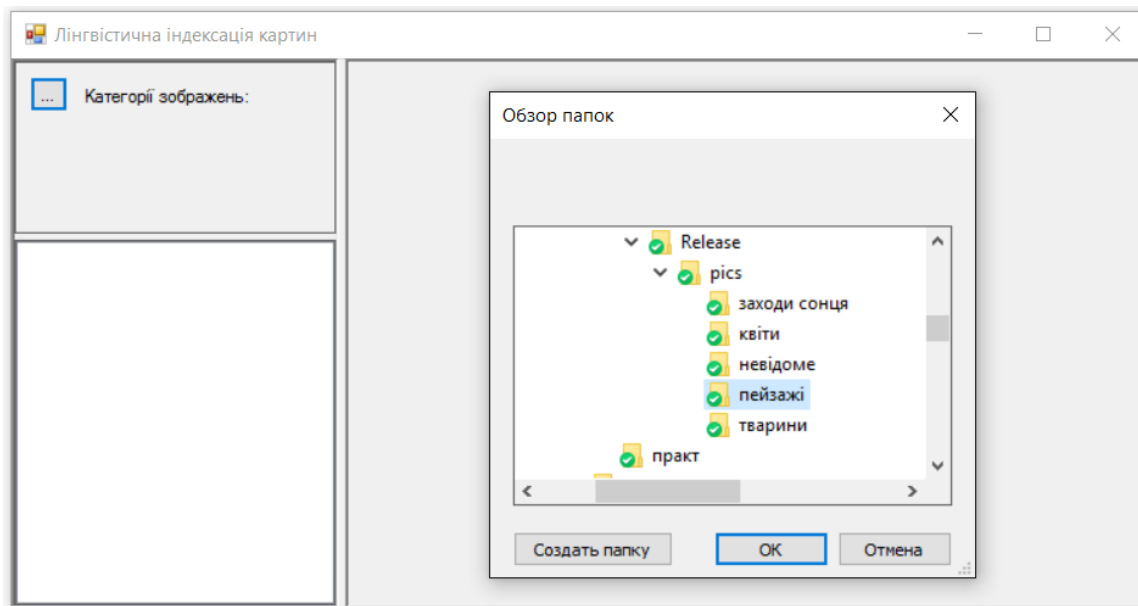
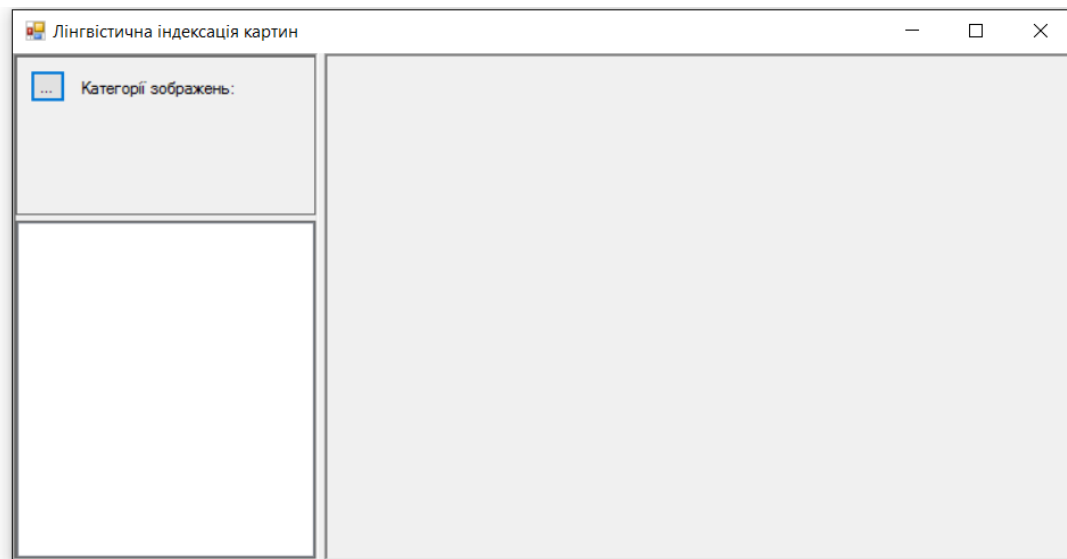
Структура проекта



Діаграма класів проекту

# Вид початкових вікон програми лінгвістичної індексації картин

Головне вікно програми

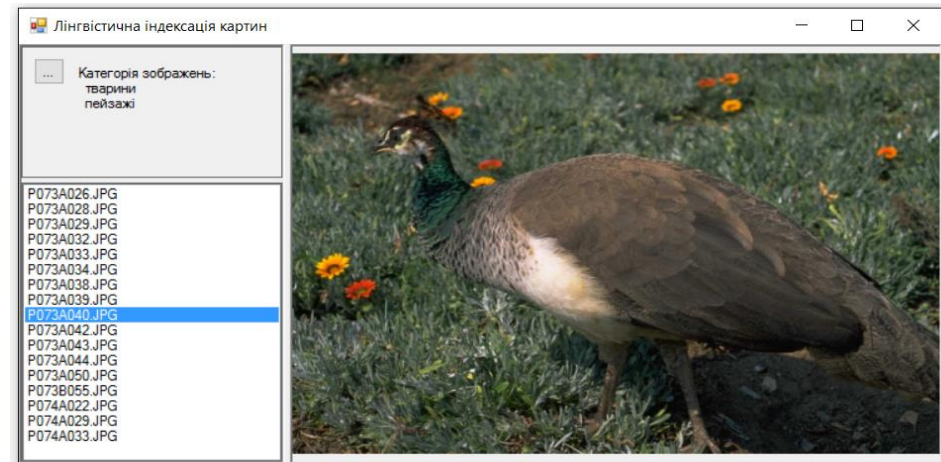
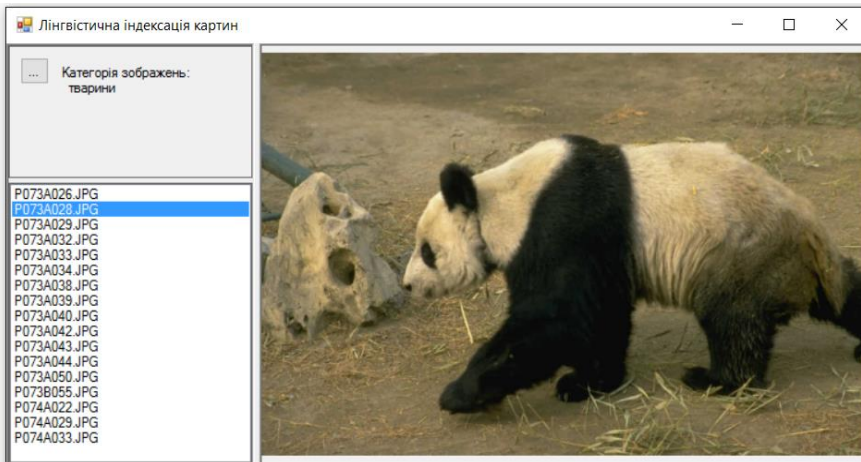


Вибір каталогу з зображеннями

# Результати роботи програми лінгвістичної індексації картин



Вікно програми з обраним каталогом



Результат роботи програми



# Результати тестування розробленої програми та програми-аналога ALIPR

Програмний засіб	Вид вибірки	Загальна к-сть категорій/зображень у всіх категоріях	Неправильно проіндексовані	Достовірність індексування, %
Програма ALIPR	Навчальна	5/500	79	84,2
	Тестова	5/500	193	61,4
	Середнє значення достовірності			72,8
Розроблена програма	Навчальна	5/500	30	94
	Тестова	5/500	115	77
	Середнє значення достовірності			85,5

Розроблена програма має вищу достовірність лінгвістичної індексації картин (85,5%), ніж аналогічна програма (72,8%), а значить мета роботи досягнута

# ЕКОНОМІЧНА ЧАСТИНА

Було здійснено економічне обґрунтування доцільності розробки інформаційної технології лінгвістичної індексації картин на основі нейронної мережі.

Встановлено, що дана програма є кращою за аналог:

- відносний показник якості складає 1,253,
- конкурентоспроможності – 1,566.

Необхідна сума інвестицій – 36460,2 грн.

Абсолютна ефективність вкладених інвестицій - 554951,2 грн ,.

Термін окупності інвестицій - менше року (8 місяців), що свідчить про економічну доцільність та ефективність фінансування розробки інформаційної технології лінгвістичної індексації картин на основі нейронної мережі

# АПРОБАЦІЯ РЕЗУЛЬТАТІВ РОБОТИ ТА ПУБЛІКАЦІЇ

## Апробація результатів роботи.

Результати роботи були апробовані на

- XLV Науково-технічній конференція ВНТУ (Україна, м. Вінниця, 2016 р.)
- XLVI Науково-технічній конференція ВНТУ (Україна, м. Вінниця, 2017 р.)
- на Десятій міжнародній науково-практичній конференції «ІНТЕРНЕТ-ОСВІТА-НАУКА-2016» (Вінниця)

## Публікації.

За результатами магістерської кваліфікаційної роботи опубліковано 3 тез доповідей на конференції.

# ВИСНОВОК

В результаті виконання роботи було розроблено інформаційну технологію лінгвістичної індексації картин на базі нейронної мережі багат шаровий персептрон. За допомогою розробленої інформаційної технології та програмних засобів користувач може проводити лінгвістичну індексацію (класифікацію по змісту) картин (зображень, поданих у цифровому форматі) . Програмну реалізацію технології здійснено об'єктно-орієнтованою мовою C#. Розроблена програма має підвищену достовірність індексації картин (85,5%) порівняно із аналогічними програмами (72,8%). Отже, мета роботи досягнута.

Дякую за увагу!