

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМ І.СІКОРСЬКОГО»
«ІНСТИТУТ ПРИКЛАДНОГО СИСТЕМНОГО АНАЛІЗУ»
КАФЕДРА МАТЕМАТИЧНИХ МЕТОДІВ СИСТЕМНОГО АНАЛІЗУ**

Апарат нейронних мереж в задачах генерування ілюстрацій на основі словесного опису

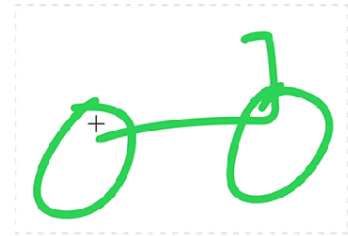
Виконав: студент групи КА-31 Логвін А.В.
Науковий керівник: к.т.н, доц. Тимощук О.Л.

Київ - 2017

Актуальність теми

1. Google редактор

Do you mean: 



2. Ілюстрації в Allo, згенеровані нейронними мережами



3. Датасет Quick, Draw!

Мета

1. Дослідження методів побудови зображень на основі словесного опису
2. Розробка програмного засобу, який реалізує генерування за допомогою нейронних мереж

Постановка задачі

1. Аналіз існуючих методів генерування зображень на основі словесного опису
2. Порівняння ефективності методів
3. Реалізація методів мовою програмування Python
4. Аналіз результатів

Об'єкт дослідження

Структури нейронних мереж, що використовуються для генерування зображень отримуючи на вхід їх словесний опис

Предмет дослідження

StackGAN та DRAW мережі

Опис вхідних даних

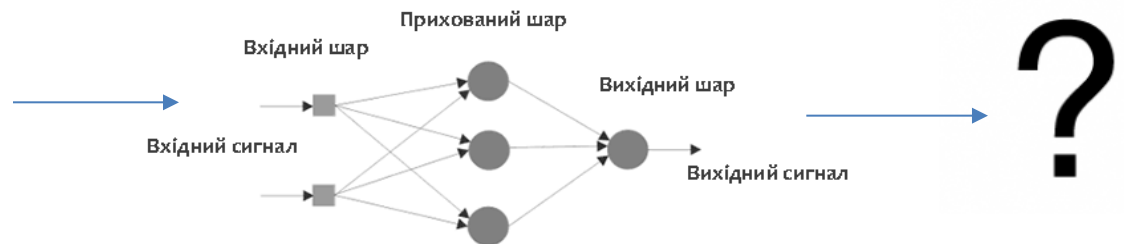
Датасет :

Цей птах має коричневі крила та жовтий живіт.

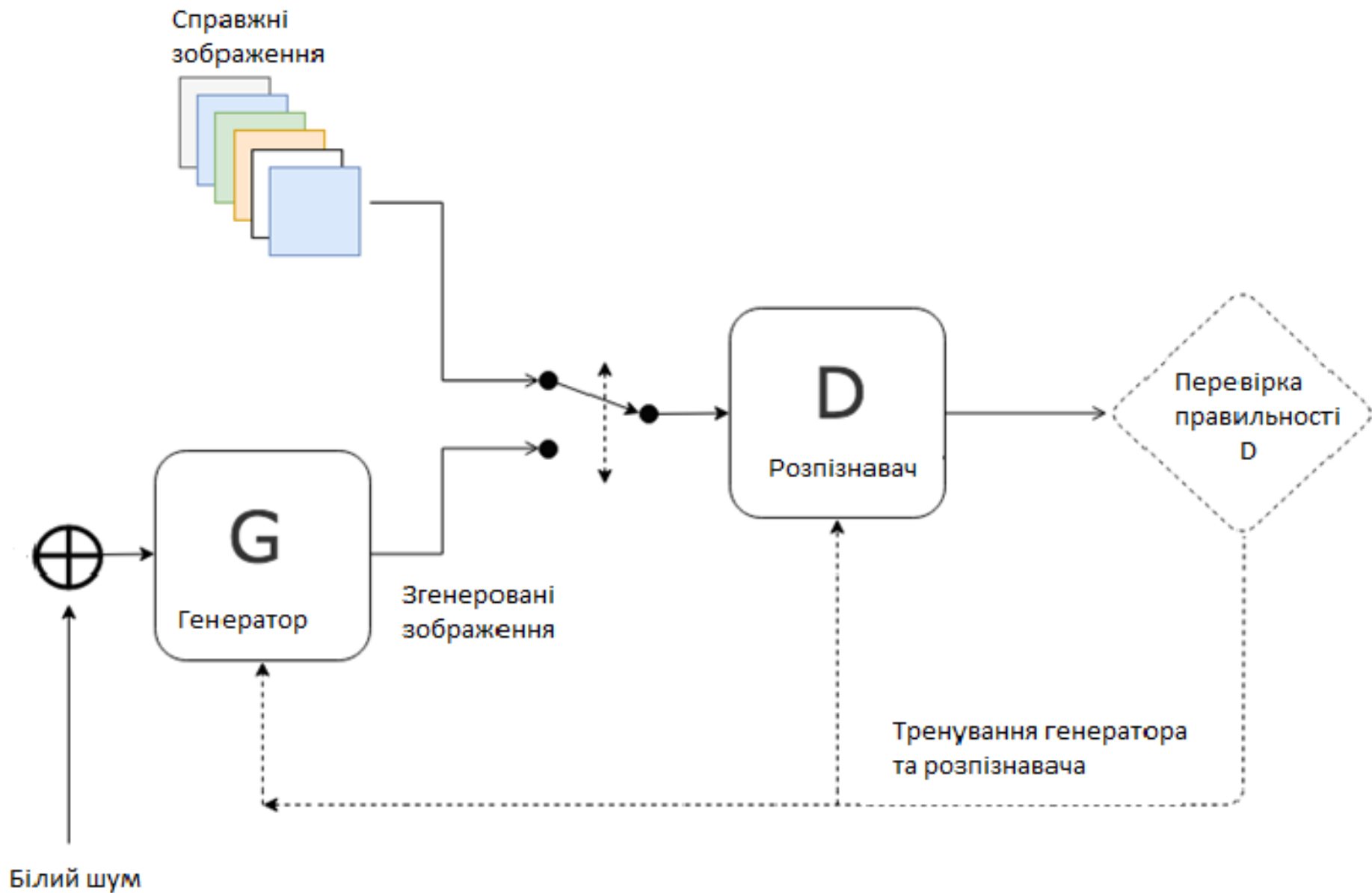


Генерується зображення з наступної інформації:

Цей птах має синьо – біле забарвлення та дуже маленький хвіст.



GAN мережі



Алгоритм роботи GAN мереж

1. $h \leftarrow \varphi(t)$ кодується відповідний опис t
2. $\hat{h} \leftarrow \varphi(\hat{t})$ кодується не відповідний опис \hat{t}
3. $z \sim N(0,1)^Z$ генерується зображення з білого шуму
4. $\hat{x} \leftarrow G(z, h)$ зображення подається на генератор, отримуємо нове зображення \hat{x}
5. $s_r \leftarrow D(x, h)$ справжнє зображення, відповідний текст
6. $s_w \leftarrow D(x, \hat{h})$ справжнє зображення, не відповідний текст
7. $s_f \leftarrow D(\hat{x}, h)$ згенероване зображення, відповідний текст
8. $L_D \leftarrow \log(s_r) + \frac{1}{2}(\log(1 - s_w) + \log(1 - s_f))$
9. $L_D \leftarrow D - \alpha \frac{\partial L_D}{\partial D}$ зміна розпізнавача, α – крок тренування
10. $L_G \leftarrow \log(s_f)$
11. $G \leftarrow G - \alpha \frac{\partial L_G}{\partial G}$ зміна генератора

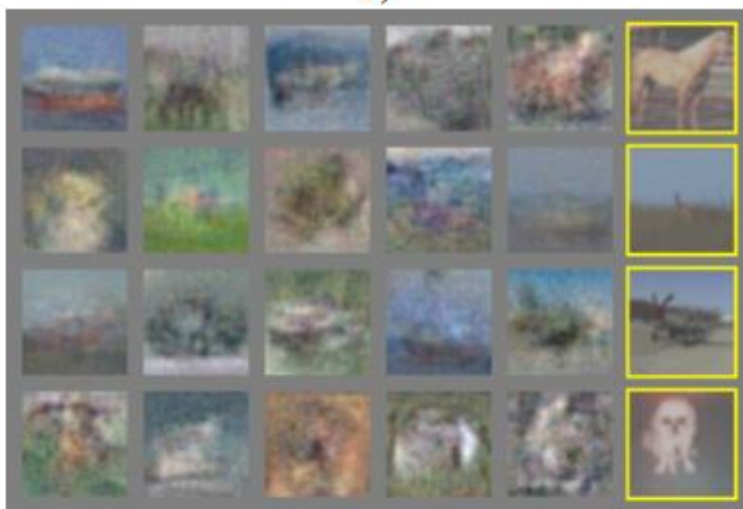
Результат роботи GAN мереж



a)



b)



c)



d)

DRAW мережі

1. $\hat{x}_t = x - \sigma(c_{t-1})$
2. $r_t = read(x_t, \hat{x}_t, h_{t-1}^{dec})$
3. $h_t^{enc} = RNN^{enc}(h_{t-1}^{enc}, [r_t, h_{t-1}^{dec}])$
4. $z_t \sim Q(Z_t | h_t^{enc})$
5. $h_t^{dec} = RNN^{dec}(h_{t-1}^{dec}, z_t)$
6. $c_t = c_{t-1} + write(h_t^{dec})$

де RNN^{enc} та RNN^{dec} - кодуєча та декодуєча мережі

x – справжнє зображення

\hat{x}_t – змінене справжнє зображення

r_t – вхідні дані для кодуєчої мережі

h_t^{enc} – закодовані вхідні дані

z_t – згенерований шум на основі частини вибірки

h_t^{dec} – декодовані дані

c_t – матриця зображення

Conditional GAN :

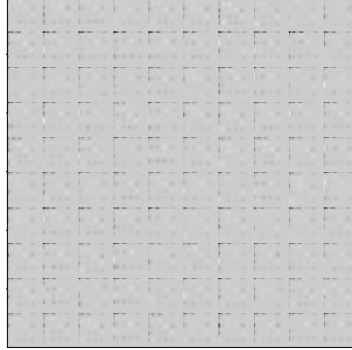
Умова (опис)

Шум

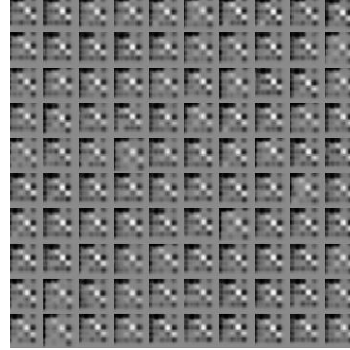
GAN

Зображення

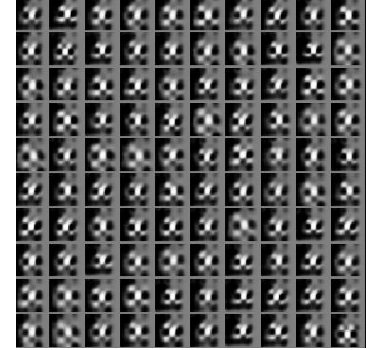
Результати програми



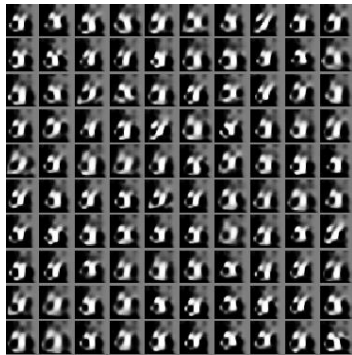
Після 1000 епох



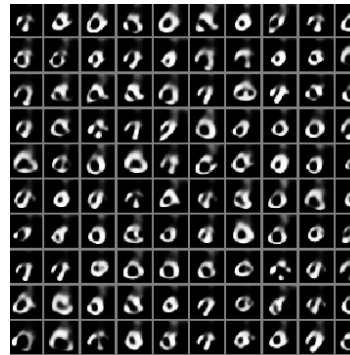
Після 2000 епох



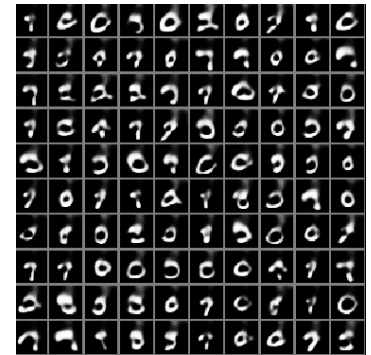
Після 3000 епох



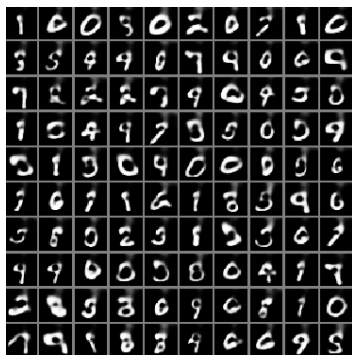
Після 4000 епох



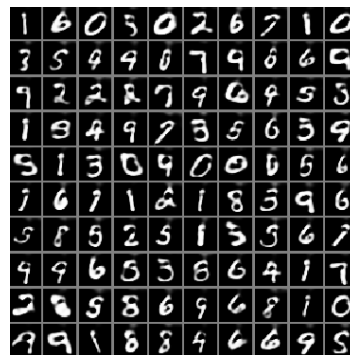
Після 5000 епох



Після 6000 епох



Після 7000 епох



Після 8000 епох



Після 9000 епох

DRAW мережі

Великий білий птах з чорними крилами та червоною головою.

This is a large **white** bird with **black wings** and a **red head**.

Малий птах з чорною головою та крилами та зеленими крилами

A small bird with a **black head and wings** and features **grey wings**.

Птах з білою грудиною, біло-коричневою головою та тонким дзьобом.

This bird has a **white breast**, brown and white coloring on its head and wings, and a thin pointy beak.

Малий птах з білим тулубом та чорними полосками на хвості, голові та крилах.

A small bird with **white base** and **black stripes** throughout its belly, head, and feathers.

Птах малого розміру з кремовим тулубом та коротким дзьобом.

A small sized bird that has a cream belly and a short pointed bill.

Цей птах повністю червоний.

This bird is **completely red**.

Цей птах повністю білий.

This bird is **completely white**.

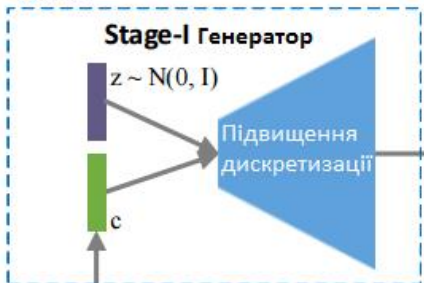
Це жовтий птах. Крила яскраво сині.

This is a **yellow** bird. The wings are **bright blue**.

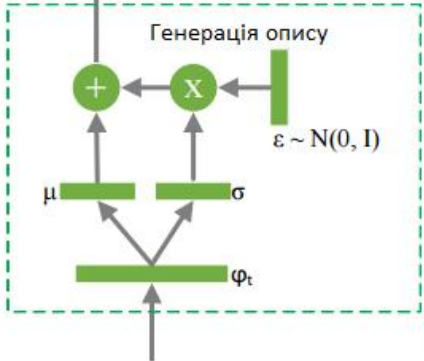
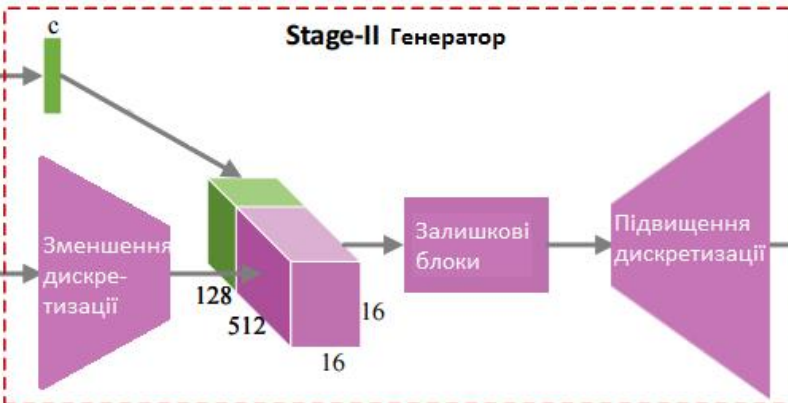


Візуалізація процесу роботи StackGAN мережі

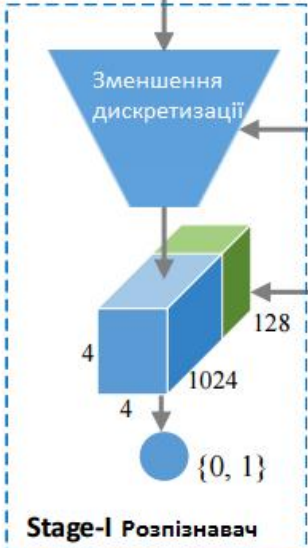
Цей птах сірого кольору з білим на грудях і має дуже короткий дзьоб.



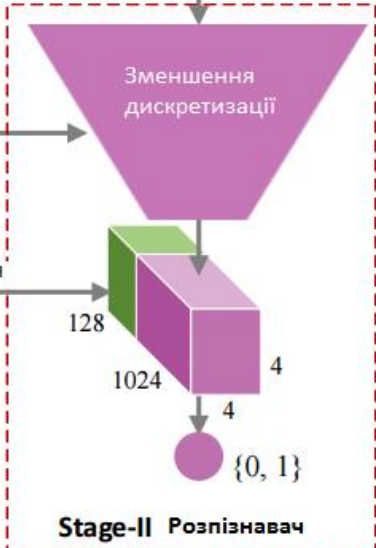
Збільшення опису



Цей птах сірого кольору з білим на грудях і має дуже короткий дзьоб.



256 x 256 з вибірки



Цей птах сірого кольору з білим на грудях і має дуже короткий дзьоб.

Результат роботи StackGAN мережі

This bird is blue with white and has a very short beak



This bird has wings that are brown and has a yellow belly



A white bird with a black crown and yellow beak



This bird is white, black, and brown in color, with a brown beak



The bird has small beak, with reddish brown crown and gray belly



This is a small, black bird with a white breast and white on the wingbars.



This bird is white black and yellow in color, with a short black beak



Аналіз роботи

- + Зображення виглядає реальним, наявні фон та деталі
- + Відповідність текстовому опису
- + Оригінальність зображення
- Необхідність різнопланової вибірки
- Низька роздільна здатність

Поради для подальших досліджень

Використання моделі на інших текстових вибірках

Повторне використання моделі для покращення роздільної здатності

Висновки

1. Проаналізовано існуючі методи генерування зображень на основі словесного опису
2. Порівняно ефективність методів
3. Реалізовано DRAW мережу мовою програмування Python
4. Виконано аналіз результатів