

# Esercitazione

Progetto

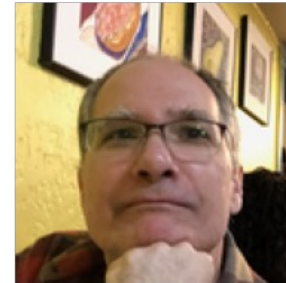
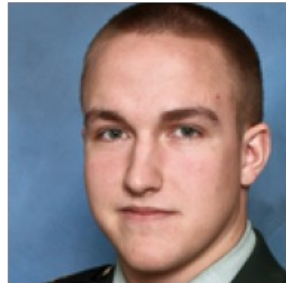
# Progetto

- Image generation
- Task: generare immagini di volti in grado di «ingannare» un modello che riconosce immagini reali e sintetiche
- Costruire un modello generativo basato sui modelli di diffusione

# Dataset

- Il progetto è costituito da 4 dataset:
  - Train: 3 dataset, ciascuno contenente 10k immagini di volti
  - Test: file di testo contenente 50 etichette. Per ogni etichetta si dovrà generare degli esempi di volto
- Ad ogni immagini è associata un etichetta che descrive alcune caratteristiche dell'immagine (es. 01100):
  - Frangia si/no
  - Occhiali si/no
  - Barba si/no
  - Sorridente si/no
  - Giovane si/no

# Dataset



Immagini 64x64

# Progetto

- Si deve costruire un modello di diffusione condizionato che data un'etichetta di classe sia in grado di generare un volto di una persona che esprime determinate caratteristiche
- L'immagine generata deve essere classificata da un modello discriminatore (preaddestrato e immutabile) come immagine reale. Il modello restituisce uno score  $[0,1]$  (1 per le immagini vere).

[https://drive.google.com/drive/folders/1mac\\_rpmnyVatFVhbcBlk6Vzope263ll0?usp=share\\_link](https://drive.google.com/drive/folders/1mac_rpmnyVatFVhbcBlk6Vzope263ll0?usp=share_link)

# Suggerimenti

- Colab/Azure/ecc. con salvataggio dello stato
- Test su porzioni del dataset per la stima degli iper-parametri
- Modelli semplici vs modelli complessi
- Loss standard vs loss custom

Info:

[https://pytorch.org/tutorials/beginner/saving\\_loading\\_models.html](https://pytorch.org/tutorials/beginner/saving_loading_models.html)

# Suggerimenti (2)

- Analisi del dataset
  - Quanti immagini utilizzare? Qual è il trade off tra numero di campioni e tempo di computazione?
- Scelta del modello
  - Quale scegliere? Studio della letteratura
  - Quali architetture?
- Analisi degli errori
- Tuning dei modelli
  - Analisi della loss

# Requisiti

- Definire e addestrare un modello di diffusione
- Descrivere in una presentazione le scelte progettuali (in particolare modello e loss utilizzate) e tutti i parametri utilizzati nella sperimentazione. Analisi di alcuni campioni generati (analisi qualitativa).
- Consegnare notebook (e/o file sorgente) utilizzati per il training, il notebook di test, presentazione e dump del modello.
- Si consegnano 10 immagini generate per ogni etichetta presente nel file di test in un archivio **test\_{MATRICOLA}.zip**



## Requisiti (2)

- Si consegnano **10** immagini generate per ogni etichetta presente nel file di test (contiene **50** label) in un archivio **test\_{MATRICOLA}.zip**
- I nomi dei file delle immagini DEVONO avere il formato **{NOME}\_{LETTERA}.jpg**
  - Il placeholder NOME corrisponde al nome associato all'etichetta presente nel file di test
    - Es. image001;10110
  - Il placeholder LETTERA è una delle prime dieci lettere dell'alfabeto per distinguere i 10 campioni per ogni etichetta
    - Es. image001\_A.jpg, image001\_B.jpg, image001\_C.jpg, ...
    - Es. image099\_A.jpg ..... image099\_J.jpg

# Protocollo di valutazione

- Valutazione generale del progetto
- Valutazione delle immagini generate dal modello sul test set.

Per ogni immagine si calcola la predizione del modello di classificazione e per ogni etichetta si seleziona lo score massimo. Il punteggio totale è dato dalla media degli score massimi

$$R = \frac{1}{n} \sum_{i=1}^n \max\{s_{i,1}, \dots, s_{i,10}\}$$

Dove s indica lo score restituito dal modello di classificazione.

Si accede all'orale se il valore R è  $\geq 0.5$  (BASELINE)

# Protocollo di valutazione

Per tutti i progetti che superano la baseline si costruisce una classifica per score decrescente.

Si assegnano 15 pt se si supera la baseline (criterio  $\geq$ ).

Si assegnano 5 pt in funzione del posizionamento nella classifica tramite la formula

$$5 * \frac{score - baseline}{\max(classifica) - baseline}$$

NB il max deve essere almeno superiore a ***baseline+15%***

# Protocollo di valutazione

Nel pacchetto di rilascio è presente uno script con cui fare la valutazione delle immagini generate.

```
python classify.py path/to/image/folder
```

# Appelli

- 30 - giugno - 9:00 - Aula MOD5
  - Consegna entro «4gg prima dell'esame». Lunedì 26 ore 12
- 24 - luglio - 9:00
- 6 - settembre - 9:00

# Proposte di tesi

## Diffusion Model

- Task di domain translation
- Generazione condizionata da testo e immagini
- Style transferring
- Etc.