



# Face Attribute Manipulation

Cristina Tîmbur, 351

Mihai-Adrian Oprea, 352



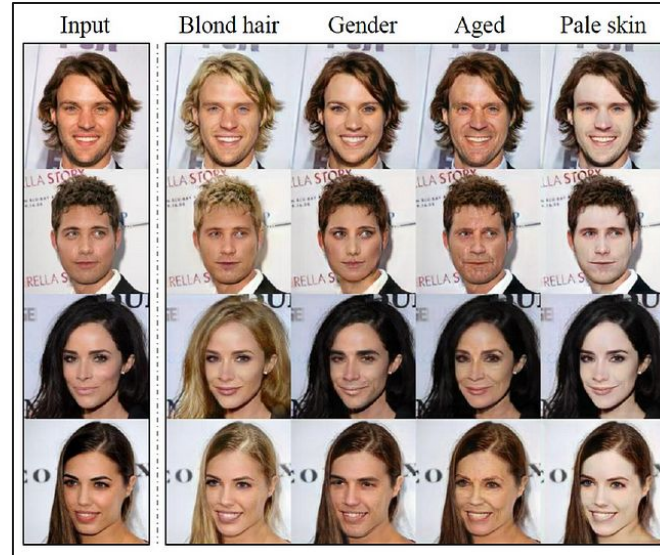
# Descrierea proiectului

01

**Tema:** Generarea imaginilor în care anumite attribute ale feței sunt schimbate (vârsta, sexul, culoarea părului, culoarea pielii etc.), pornind de la o imagine de input.

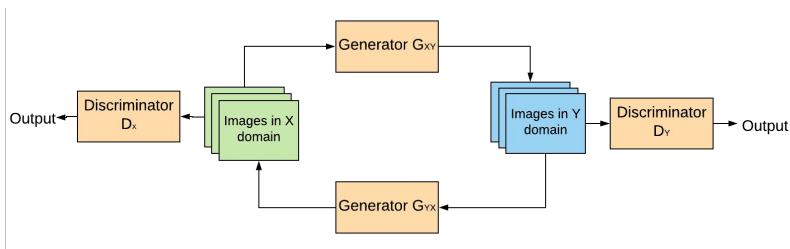
02

**Abordarea aleasă:** StarGAN (Choi et al.) – o rețea generativ adversarială care învață transferul unor trăsături din mai multe domenii.



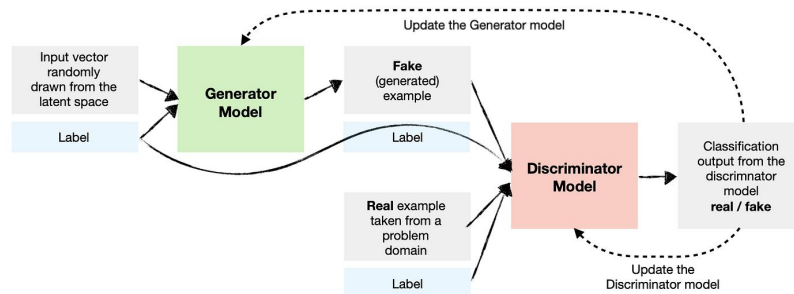
# Arhitectura StarGAN

## CycleGAN



- ❑ Învățăm transformări reversibile (după ce modificăm un atribut facial, facem și transformarea inversă; obiectivul este să reconstruim imaginea inițială).
- ❑ În cazul modelului StarGAN, deoarece putem condiționa generarea după etichete, atât transformarea directă cât și cea inversă sunt făcute de același generator.

## ConditionalGAN



- ❑ Generăm imagini care au anumite atribute, specificate printr-o etichetă. Imaginile generate trebuie să fie realiste și să corespundă atributelor specificate.
- ❑ **Loss discriminator** = Loss adversarial + Loss clasificare
- ❑ **Loss generator** = Loss adversarial + Loss clasificare + Loss reconstrucție

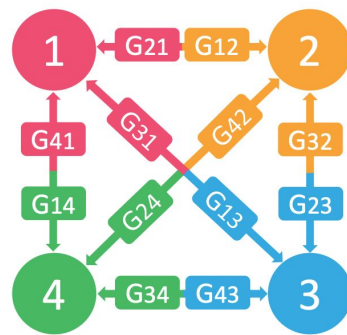
# Proprietăți StarGAN

**Abordare clasică:** Antrenăm câte un GAN pentru fiecare schimbare de atribut pe care o dorim.

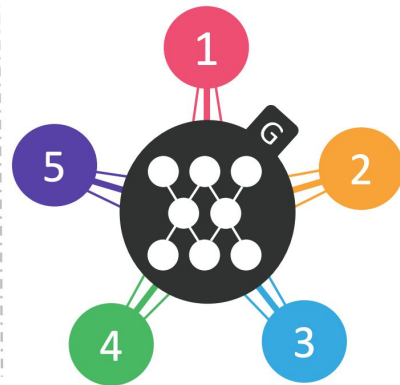
**StarGAN:** Antrenăm un singur model, eventual pe seturi de date neomogene, care să poată realiza orice schimbare de atribut pe care o dorim.

**Seturi de date neomogene:** Putem antrena un model pe seturi de date care conțin imagini cu atribute diferite construind etichete compuse pentru imagini.

(a) Cross-domain models



(b) StarGAN



# Rezultatele proiectului

01

## **Extinderea codului open-source StarGAN**

Extensia noastră permite antrenarea modelului pe un număr arbitrar de seturi de date neomogene.

Codul sursă este publicat la acest [link](#).

02

## **Antrenarea StarGAN pe 3 seturi de date**

Am antrenat un model StarGAN care poate modifica 9 atribute ale portretelor:

- Zâmbetul
- Sprâncenele
- Stilul (comic sketch)
- Culoarea părului
  - Blond
  - Brunet
  - Negru
- Sexul
- Vârsta
- Bretonul

# Generic Dataset

## CelebA Format Datasets

Seturile de date care au atributele binare adnotate pentru fiecare imagine: genul, culoarea părului, dacă persoana poartă ochelari etc.

Constructor: calea către un director cu imagini și un fișier care descrie atributele binare prezente în fiecare imagine.



## RaFD Format Datasets

Seturile de date care grupează persoanele în categorii total disjuncte, de exemplu: persoane care zâmbesc, persoane care plâng etc.

Constructor: calea către un director care conține mai multe subdirectoare corespunzătoare fiecărei categorii de imagini.

# Achiziția datelor

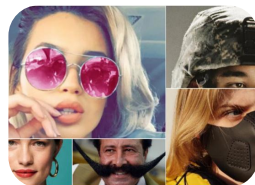
## CelebA



Conține 200k de imagini cu adnotări binare pentru 40 de atribute de interes

Dimensiune: 178 x 218

## Face Attributes Grouped



Am folosit 1000 de imagini cu oameni care au mustață sau care poartă ochelari de soare

Dimensiune: 512 x 512

## Comic Faces



Conține 10k de perechi de poze originale și comic sketch-ul asociat

Dimensiune: 1024 x 1024

Redimensionarea imaginilor la 178 x 218:

- Detecția fețelor (centrare) + redimensionare
- Adăugare padding + redimensionare

# Custom solver

## Clasa originală Solver

- ✓ numărul de straturi convoluționale din G și D
- ✓ rata de învățare
- ✓ metode pentru antrenarea pe setul de date CelebA, pe RaFD sau pe ambele
- ✗ metode pentru antrenarea pe un număr arbitrar de seturi de date



## Clasa noastră CustomSolver

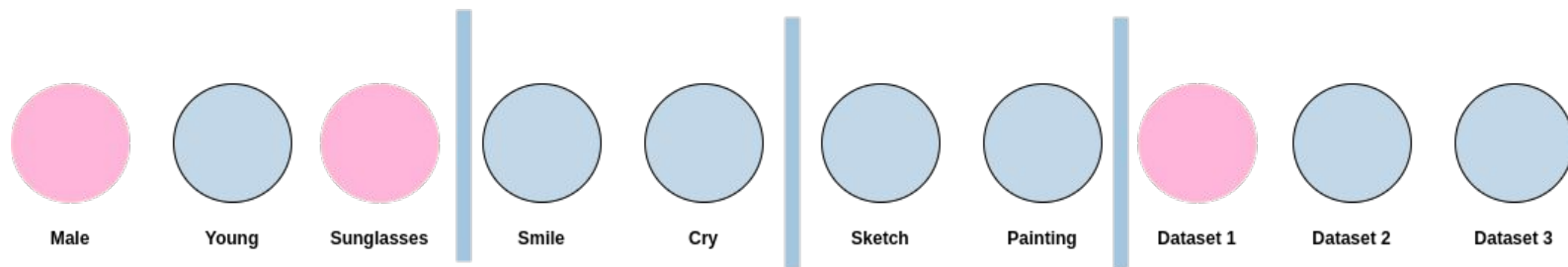
- ✓ număr arbitrar de Generic Datasets
- ✓ generează etichete corecte pentru antrenare
- ✓ antrenează pe rând câte un batch din fiecare set de date
- ✓ păstrează unele facilități ale clasei originale (checkpointing periodic, scăderea lr-ului etc.)



# Generarea etichetelor

**Eticheta** este formată din două părți:

- ❑ Concatenarea tuturor atributelor din toate seturile de date
- ❑ O mască care indică ce set de date este folosit



Exemplu de construcție a etichetei în cazul antrenării pe 3 seturi de date, când imaginea curentă este a unui bărbat în vârstă, care poartă ochelari de soare.

# Antrenarea modelului

## Un singur dataset

- ❑ CelebA cu 5 attribute selectate
- ❑ 20k iterații
- ❑ rezultate comparabile calitativ cu modelul StarGAN public

## Două dataset-uri

- ❑ CelebA cu 5 attribute selectate
- ❑ Face Attributes Grouped cu 2 categorii selectate
- ❑ 20k iterații
- ❑ rezultate slabe din cauza categoriilor nu neapărat disjuncte

## Trei dataset-uri

- ❑ CelebA cu 5 attribute selectate
- ❑ Celeba cu alte 3 attribute selectate
- ❑ Comic faces
- ❑ 75k iterații
- ❑ rezultate bune pentru toate attributele învățate

Observații antrenare:

- Rata de învățare a fost scăzută liniar după un anumit număr de iterații
- Durata antrenării pe 3 seturi de date a fost de aproximativ **40 de ore** pe GPU P100 Kaggle

# Rezultate pe imaginile de antrenare



Original | Sex | Vârsta | Negru | Blond | Şaten | Zâmbet | Brows | Breton | Comic

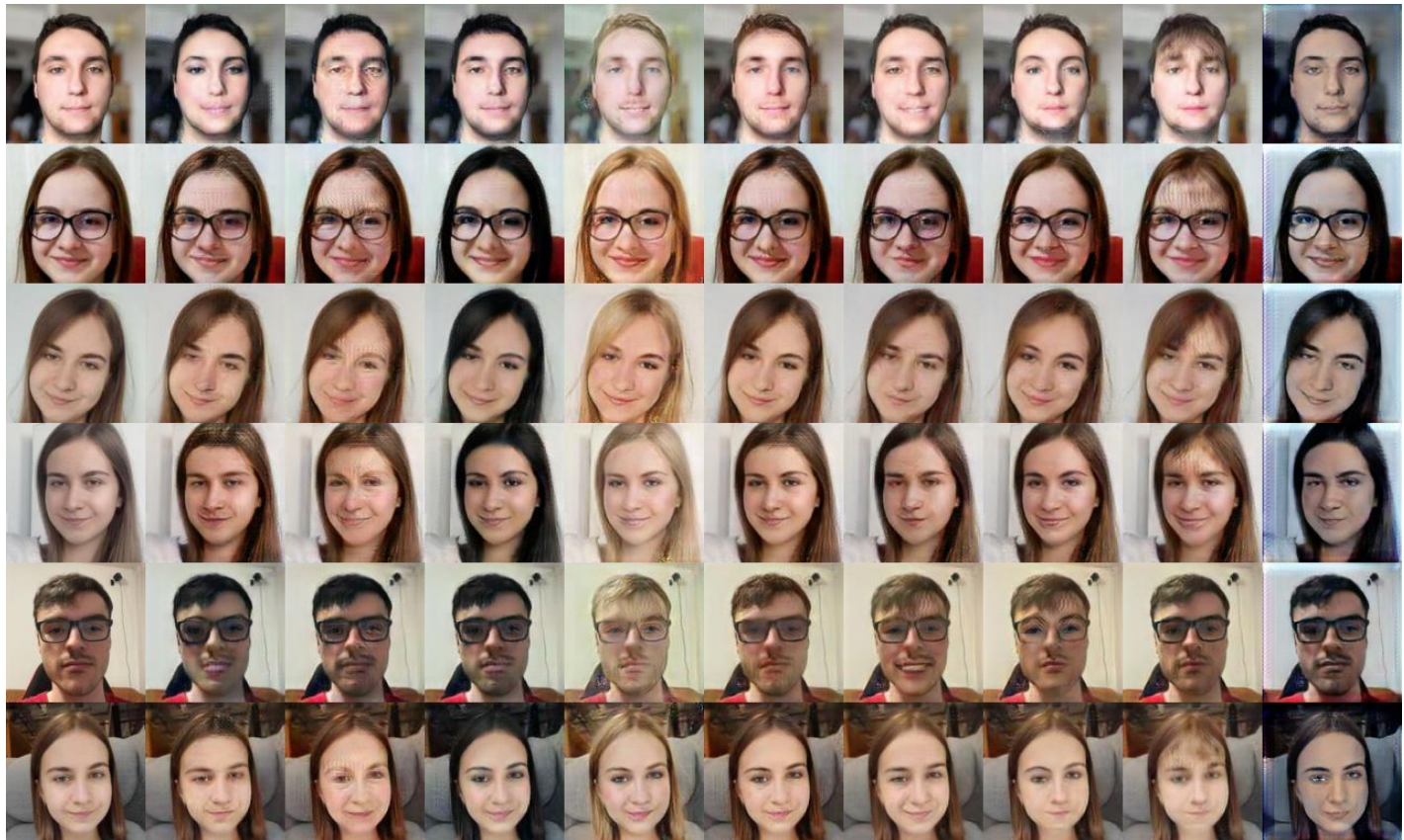


# Rezultate pe imaginile de antrenare (2)



Original | Sex | Vârsta | Negru | Blond | Şaten | Zâmbet | Brows | Breton | Comic

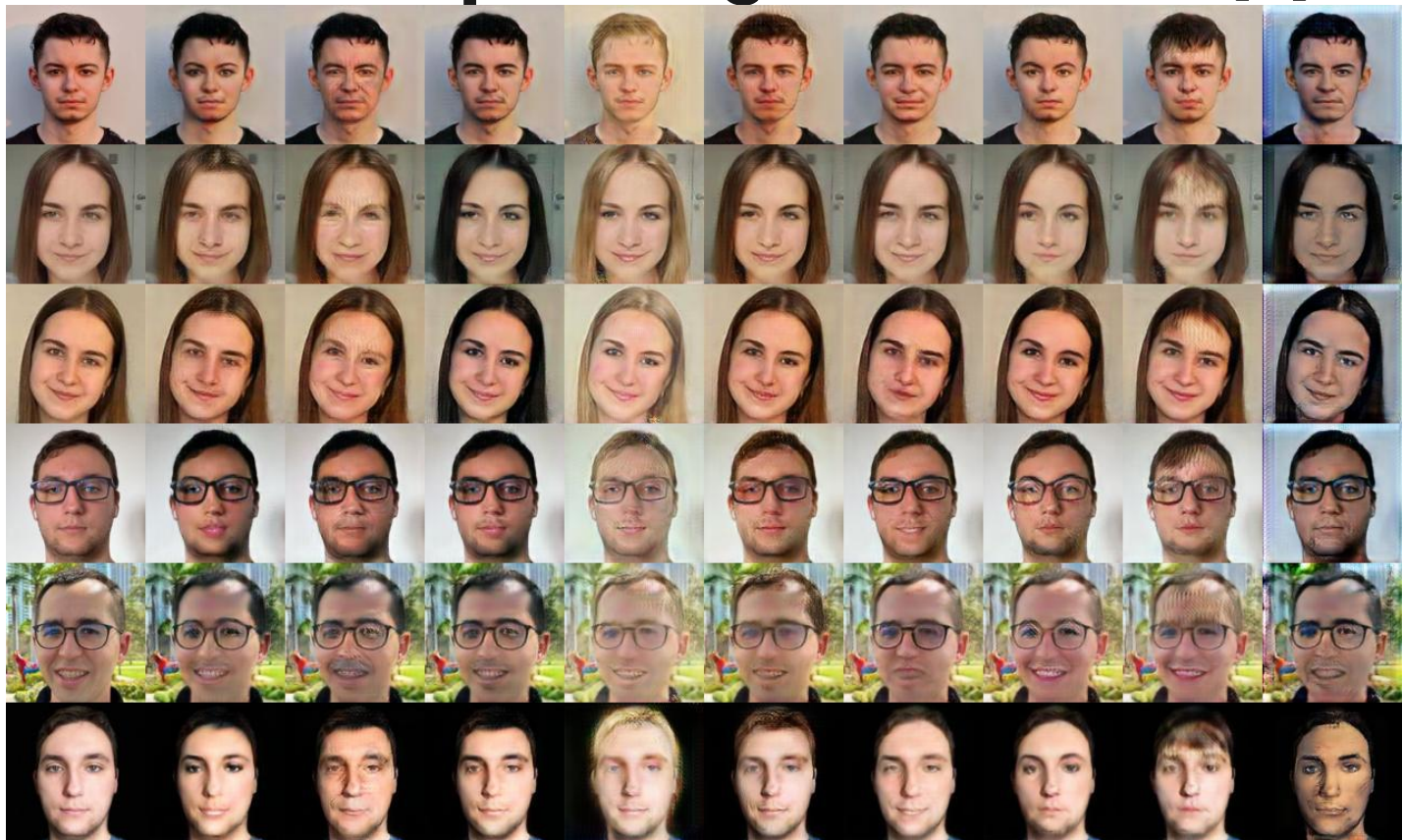
# Rezultate pe imaginile de test



Original | Sex | Vârsta | Negru | Blond | Şaten | Zâmbet | Brows | Breton | Comic



# Rezultate pe imaginile de test (2)



Original | Sex | Vârstă | Negru | Blond | Şaten | Zâmbet | Brows | Breton | Comic

# Interpretarea rezultatelor

Modelul a învățat bine unele transformări subtile (cum ar fi schimbare de gen, zâmbet), dar nu se descurcă la fel de bine la adăugarea sau scoaterea bretonului, la persoane care poartă ochelari sau la variații mari de luminozitate în imagine.

## Direcții de **îmbunătățire a rezultatelor**:

- ❑ Augmentare mai diversă a datelor (schimbări de contrast, de luminozitate etc.)
- ❑ Mai multe date (CelebA are 200k imagini, Comic Faces are 20k)
- ❑ Mai mult antrenament (40 ore pentru modelul cu 3 dataset-uri vs. 1 săptămână pentru modelul StarGAN public)

**Vă mulțumim!**