Department of Computer, Control, and Management Engineering “Antonio Ruberti”

4th January, 2023

DDIM-based Synthetic Medical Image Augmentation for increased CNN Performance in COVID-19 Classification

Generative Deep Learning A.Y. 2021/2022 - PhD Engineering in Computer Science

Student: Simone Rossetti
Email: rossetti@diag.uniroma1.it

PhD tutor: Prof.ssa Fiora Pirri
Email: pirri@diag.uniroma1.it

SAPIENZA
UNIVERSITÀ DI ROMA

Outline

- Introduction
  - Synthetic Data Augmentation
- Background
  - Diffusion Models
  - Score Based Models
  - DDPMs, DDIMs
- Method
  - DDIM Synthetic Augmentation
  - Interpolation and Inpainting
- Results
- Conclusions

Introduction

Synthetic Data Augmentation

Problem: Small and weak datasets are a recurrent problem in the medical imaging domain.

Causes:

- Few available datasets, most of them are limited in size and only applicable to specific medical problems;
- Collecting medical data is a complex and expensive procedure for which only experts in the field are capable to;
- ...

General solution: researchers attempt to overcome this problem by using data augmentation.

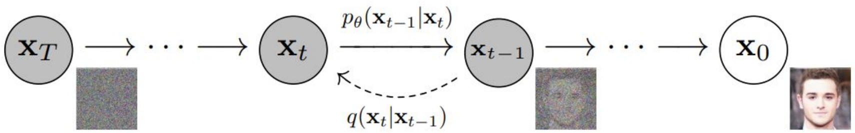
Our goals:

- implementation of a new research approach [**fri18**] and application to a different dataset [**coh20**];
- high quality synthesis of images using *denoising diffusion implicit models (DDIMs)* [**son21**];
- high classification improvements by the usage of synthetic augmentation w.r.t classic augmentation;
- ...

Set/Class	Covid	Normal	Pneumonia_bac	Pneumonia_vir	Total
Train	60	70	70	70	270
Test	9	9	9	9	36
Total	69	79	79	79	306

Background

Diffusion Models [soh15]



- DMs are a family of flexible and tractable Markov chain generative models derived from Langevin dynamics and Annealed Importance Sampling (AIS)

$$p_\theta(\mathbf{x}_0) := \int p(\mathbf{x}_{0:T}) d\mathbf{x}_{1:T}$$

- It is defined as a Markov chain with learned Gaussian transitions starting by pure noise (learns to iteratively denoise the signal)

$$p(\mathbf{x}_{0:T}) := p(\mathbf{x}_T) \prod_{i=1}^{T} p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t), \qquad p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t) := \mathcal{N}(\mathbf{x}_{t-1}; \boldsymbol{\mu}_\theta(\mathbf{x}_t, t), \boldsymbol{\Sigma}_\theta(\mathbf{x}_t, t)) \qquad p(\mathbf{x}_T) = \mathcal{N}(\mathbf{x}_T; \mathbf{0}, \mathbf{I})$$

- the forward diffusion process is a high-dimensional fixed Markov chain starting from natural images (iteratively degrades the signal)

$$q(\mathbf{x}_{1:T}|\mathbf{x}_0) = \prod_{i=1}^{T} q(\mathbf{x}_t|\mathbf{x}_{t-1}), \qquad q(\mathbf{x}_t|\mathbf{x}_{t-1}) := \mathcal{N}(\mathbf{x}_t; \sqrt{1-\beta_t}\mathbf{x}_t, \beta_t \mathbf{I})$$

- Optimize the variational bound on negative log likelihood

$$\mathbb{E}_q \left[-\log p(\mathbf{x}_T) - \sum_{t=1}^{T} \log \frac{p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)}{q(\mathbf{x}_t|\mathbf{x}_{t-1})} \right] =: L$$

DDPMs, DDIMs [ho20, son21]

$$q(\mathbf{x}_t|\mathbf{x}_0) = \mathcal{N}(\mathbf{x}_t; \sqrt{\alpha_t}\mathbf{x}_0, (1-\alpha_t)\mathbf{I}), \text{ where } \alpha_t := \prod_{s=1}^{t} 1-\beta_s$$

$$\mathbf{x}_t = \sqrt{\alpha_t}\mathbf{x}_0 + \sqrt{1-\alpha_t}\epsilon \qquad \epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$$

- DDPMs accelerate training
  - Optimize random terms of diffusion thanks to gaussian transitions chain property
  - Equivalent parametrization to noise regression

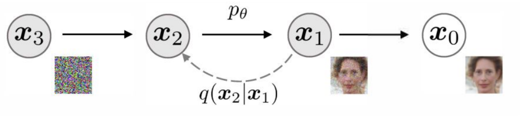
$$\mathbb{E}_q\left[\underbrace{D_{KL}(q(\mathbf{x}_T|\mathbf{x}_0) \parallel p(\mathbf{x}_T))}_{L_T} + \sum_{t=2}^{T}\underbrace{D_{KL}(q(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) \parallel p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t))}_{L_{t-1}} - \underbrace{\log p_\theta(\mathbf{x}_0|\mathbf{x}_1)}_{L_0}\right] \approx \mathbb{E}_{\mathbf{x}_0,\epsilon}\left[\parallel \epsilon - \epsilon_\theta(\sqrt{\alpha_t}\mathbf{x}_0 + \sqrt{1-\alpha_t}\epsilon, t) \parallel^2\right]$$

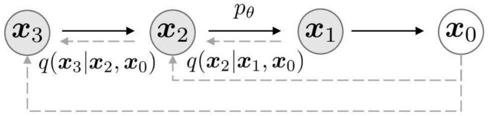
  - Sampling requires developing the full chain

$$\mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(\mathbf{x}_t - \frac{\beta_t}{\sqrt{1-\alpha_t}}\epsilon_\theta(\mathbf{x}_t, t)\right) + \sigma_t\mathbf{z} \qquad \mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$$

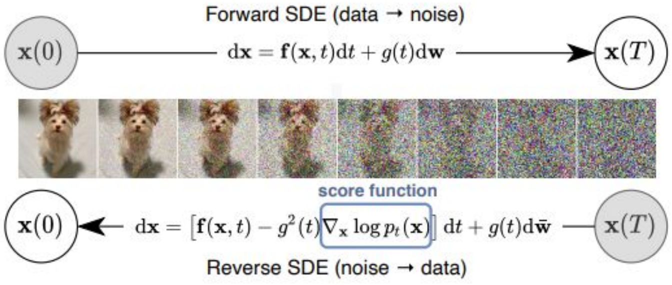
- DDIMs accelerate sampling
  - Reverse process is non-Markovian
  - Deterministic process for sigma=0

$$\mathbf{x}_{t-1} = \sqrt{\alpha_{t-1}}\underbrace{\left(\frac{\mathbf{x}_t - \sqrt{1-\alpha_t}\epsilon_\theta^{(t)}(\mathbf{x}_t)}{\sqrt{\alpha_t}}\right)}_{\text{predicted } \mathbf{x}_0} + \underbrace{\sqrt{1-\alpha_{t-1}-\sigma_t^2}\,\epsilon_\theta^{(t)}(\mathbf{x}_t)}_{\text{direction pointing to } \mathbf{x}_t} + \underbrace{\sigma_t\epsilon_t}_{\text{random noise}}$$





Score Based Models [son19]



- Try to estimate the score (gradient vector field) of a distribution

$$\frac{1}{2}\mathbb{E}_{q_\sigma(\tilde{\mathbf{x}}|\mathbf{x})p_{\text{data}}(\mathbf{x})}[\|\mathbf{s}_{\boldsymbol{\theta}}(\tilde{\mathbf{x}}) - \nabla_{\tilde{\mathbf{x}}} \log q_\sigma(\tilde{\mathbf{x}} \mid \mathbf{x})\|_2^2]$$

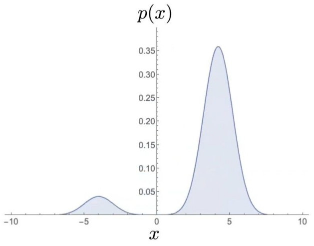
- Sampling is crossing the vector field to reach equilibrium point (Langevin Dynamics)

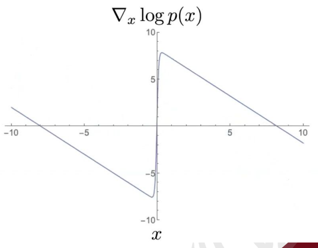
$$\tilde{\mathbf{x}}_t = \tilde{\mathbf{x}}_{t-1} + \frac{\epsilon}{2}\nabla_{\mathbf{x}} \log p(\tilde{\mathbf{x}}_{t-1}) + \sqrt{\epsilon}\, \mathbf{z}_t, \qquad \mathbf{z}_t \sim \mathcal{N}(0, I)$$

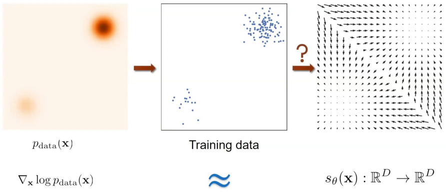
- Escape low density regions perturbing the data with random Gaussian noise of various magnitudes (Annealed Langevin Dynamics)

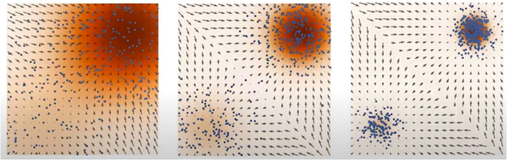
$f_\theta(\mathbf{x}) \in \mathbb{R}$ $\qquad \mathbb{E}_{p_{\text{data}}}[-\log p_\theta(\mathbf{x})] = \mathbb{E}_{p_{\text{data}}}[\log f_\theta(\mathbf{x}) - \cancel{\log Z_\theta}]$

$p_\theta(\mathbf{x}) = \dfrac{e^{-f_\theta(\mathbf{x})}}{Z_\theta}$ $\qquad \nabla_{\mathbf{x}} \log p_\theta(\mathbf{x}) = -\nabla_{\mathbf{x}} f_\theta(x) - \nabla_{\mathbf{x}} \log Z_\theta$



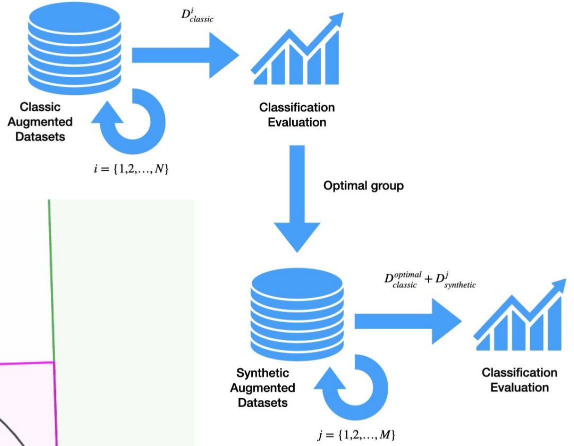






Method

DDIM Synthetic Augmentation

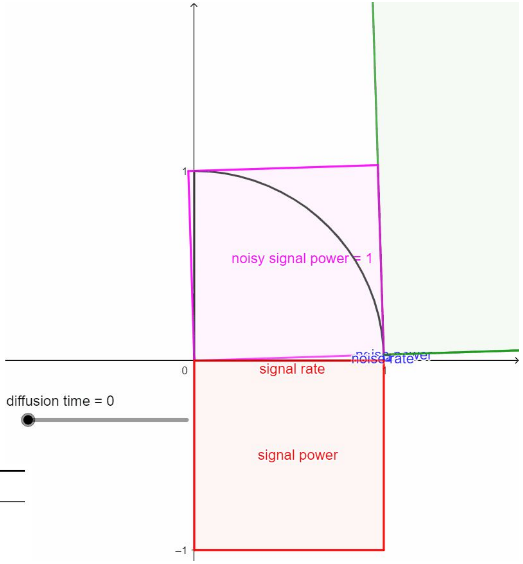


Classification task:

- We follow [**ros20**];
- We test a CNN on classical augmented sets and compare wrt fine-tuning on synthetic augmented sets;

Generative task:

- We follow https://keras.io/examples/generative/ddim/
- We choose a small U-Net as denoising network;
- Cosine schedule for alpha coefficient;



Algorithm 1 Training

repeat
  $\mathbf{x}_0 \sim q(\mathbf{x}_0)$
  $t \sim \mathrm{Uniform}(\{1,\ldots,T\})$
  $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
  Take gradient descent step on
    $\nabla_\theta \left\| \boldsymbol{\epsilon} - \boldsymbol{\epsilon}_\theta(\sqrt{\bar\alpha_t}\mathbf{x}_0 + \sqrt{1-\bar\alpha_t}\boldsymbol{\epsilon}, t) \right\|^2$
until converged

Algorithm 2 Sampling

$\mathbf{x}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
for $t = T, \ldots, 1$ **do**
  $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ if $t > 1$, else $\mathbf{z} = \mathbf{0}$
  $\mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(\mathbf{x}_t - \frac{1-\alpha_t}{\sqrt{1-\bar\alpha_t}}\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)\right) + \sigma_t \mathbf{z}$
end for
return $\mathbf{x}_0$

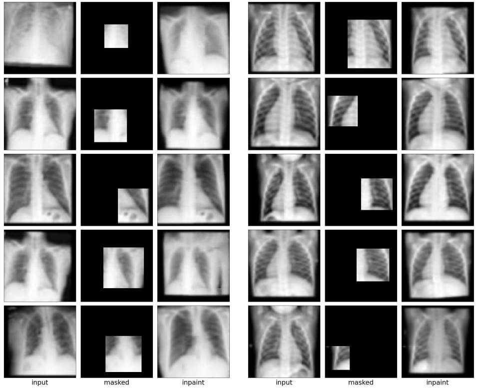
Interpolation and Inpainting

- DMs allow for inpainting (easy conditioning)

$$\mathbf{x}_{t-1}^{keep} = \sqrt{\alpha_{t-1}}\mathbf{x}_0 + \sqrt{1-\alpha_{t-1}-\sigma_t^2}\epsilon_\theta^{(t)}(\mathbf{x}_t) + \sigma_t\epsilon_t$$

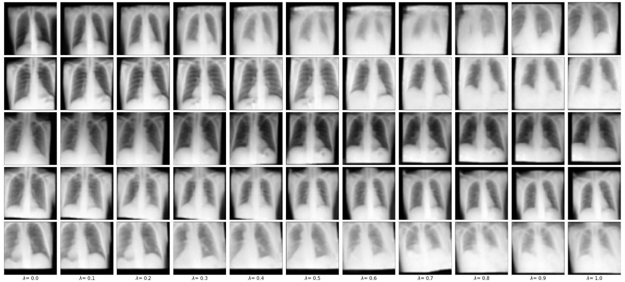
$$\mathbf{x}_{t-1}^{replace} = \sqrt{\alpha_{t-1}}\left(\frac{\mathbf{x}_t - \sqrt{1-\alpha_t}\epsilon_\theta^{(t)}(\mathbf{x}_t)}{\sqrt{\alpha_t}}\right) + \sqrt{1-\alpha_{t-1}-\sigma_t^2}\epsilon_\theta^{(t)}(\mathbf{x}_t) + \sigma_t\epsilon_t$$

$$\mathbf{x}_{t-1} = \mathbf{m} * \mathbf{x}_{t-1}^{keep} + (\mathbf{1}-\mathbf{m}) * \mathbf{x}_{t-1}^{replace}$$



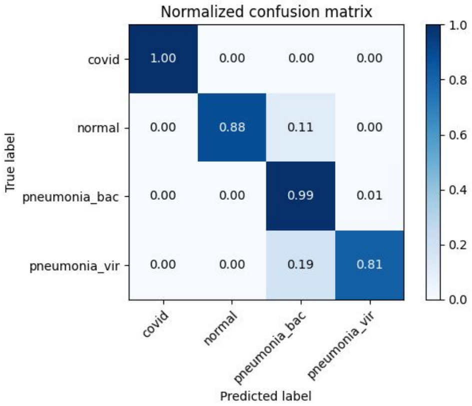
- DDIMs allow spherical interpolation in latent space

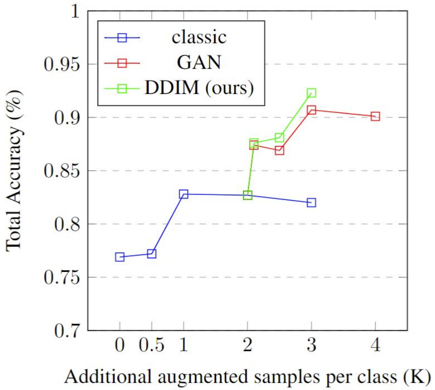
$$\mathbf{x}_T^\lambda = \frac{\sin((1-\lambda)\theta)}{\sin(\theta)}\mathbf{x}_T^0 + \frac{\sin(\lambda\theta)}{\sin(\theta)}\mathbf{x}_T^1, \quad \text{with} \quad \theta = \arccos\left(\frac{(\mathbf{x}_T^0)^\top \mathbf{x}_T^1}{\| \mathbf{x}_T^0 \| \| \mathbf{x}_T^1 \|}\right)$$



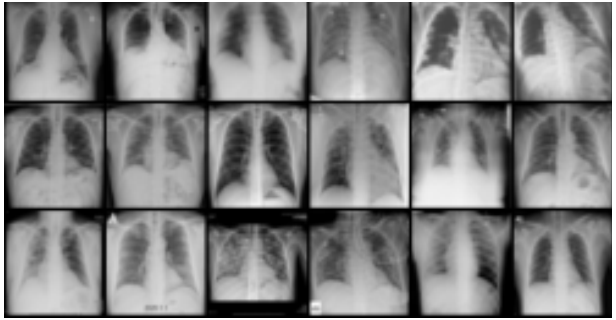
covid
DDIM

Results





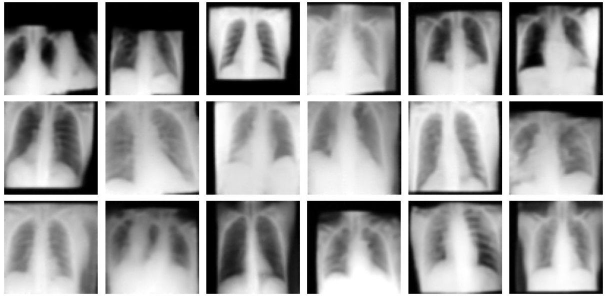
covid
real



4-fold_720_ep_0.907_acc	Accuracy	Sensitivity	Specificity	F1Score
covid	1	1	1	1
normal	0.972	0.896	0.997	0.943
pneumonia_bac	0.908	0.907	0.909	0.907
pneumonia_vir	0.933	0.824	0.97	0.888

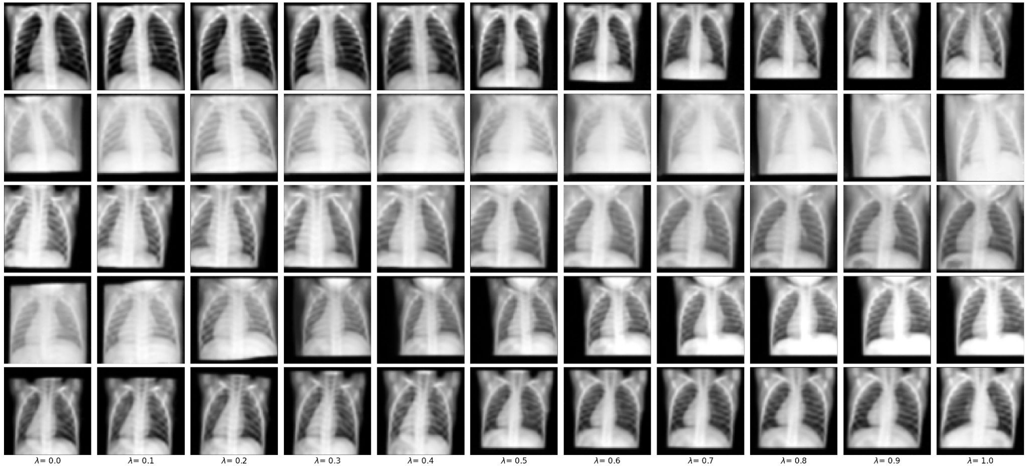
GAN
[ros20]

covid
DDIM

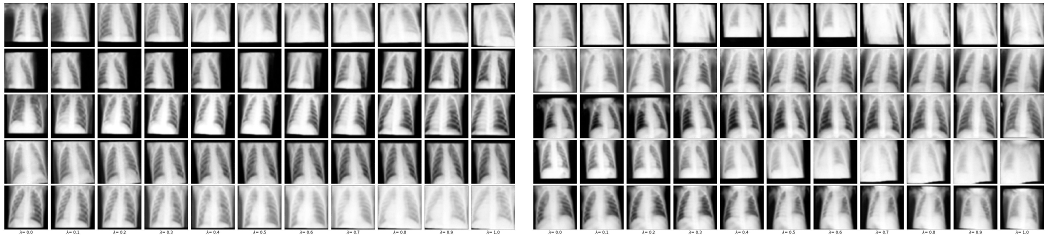


4-fold_350_ep_0.923_acc	Accuracy	Sensitivity	Specificity	F1Score
covid	1	1	1	1
normal	0.971	0.884	1	0.939
pneumonia_bac	0.923	0.993	0.9	0.944
pneumonia_vir	0.951	0.813	0.997	0.894

Results

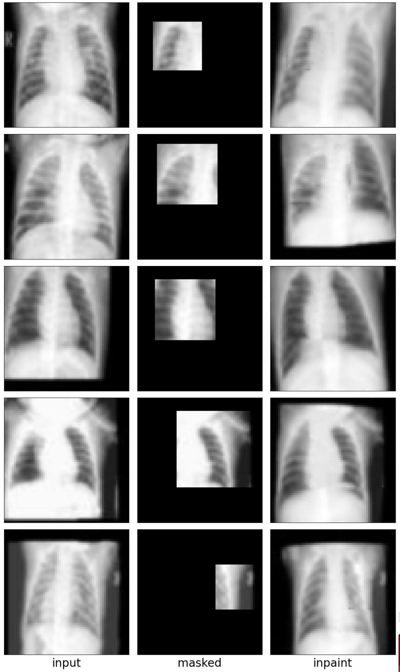


normal
DDIM



Pneumonia
bac (l), vir (r)
DDIM

Pneumonia
vir
DDIM



Conclusions

- Lack of data can be mitigated by generative models;
- Diffusion models are competitive in image quality and synthetic augmentation;
- Diffusion models easily allow for *inpainting* and *interpolation* in latent space, enhancement of synthetic augmentation;
- A *class conditioned generator* could lead to better samples?

References

[**ros20**] Simone Rossetti and Akash Garg. “GAN-based Synthetic Medical Image Augmentation for increased CNN Performance in COVID-19 Classification”. 2020. URL: https://github.com/rossettisimone/AUGMENTATION_GAN

[**ho20**] Jonathan Ho, Ajay Jain, and Pieter Abbeel. “Denoising Diffusion Probabilistic Models”. 2020.

[**soh15**] Jascha Sohl-Dickstein et al. “Deep Unsupervised Learning using Nonequilibrium Thermodynamics”. 2015.

[**son21**] Jiaming Song, Chenlin Meng, and Stefano Ermon. “Denoising Diffusion Implicit Models”. 2021.

[**son19**] Yang Song and Stefano Ermon. “Generative Modeling by Estimating Gradients of the Data Distribution”. 2019.

[**coh20**] Joseph Paul Cohen, Paul Morrison, and Lan Dao. “COVID-19 image data collection”. 2020. URL: https://github.com/ieee8023/covid- chestxray-dataset

[**fri18**] Maayan Frid-Adar et al. “GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification”. 2018. URL: https://doi.org/10.1016%2Fj.neucom.2018.09.013

Department of Computer, Control, and Management Engineering “Antonio Ruberti”

4th January, 2023

DDIM-based Synthetic Medical Image Augmentation for increased CNN Performance in COVID-19 Classification

Generative Deep Learning A.Y. 2021/2022 - PhD Engineering in Computer Science

Student: Simone Rossetti | Email: rossetti@diag.uniroma1.it

PhD tutor: Prof.ssa Fiora Pirri | Email: pirri@diag.uniroma1.it

SAPIENZA
UNIVERSITÀ DI ROMA