Enhancing Synchrotron Low-Dose Computed Tomography Image Quality Using Diffusion-Based Generative Models

Candidate: Paulo Baraldi Mausbach

Supervisor: Prof. Dr. Zanoni Dias

Co-supervisor: Prof. Dr. Hélio Pedrini

Master’s Degree Proposal

Institute of Computing - University of Campinas

October 17th of 2024

Presentation Outline

- Introduction
  - Motivation
  - Problem Definition
  - Research Questions
  - Goals and Expected Results
- Literature Review
  - Physics Concepts
  - Computation Concepts
- Materials and Methods
  - Datasets
  - Metrics
  - Methodology
- Work Plan and Execution Timeline

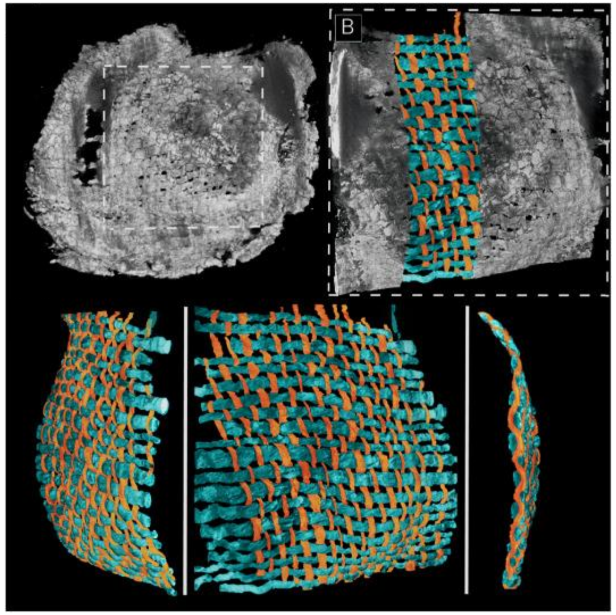
- X-Ray Computed Tomography (CT) offers a **non-invasive** technique for assessing internal structures of objects

- Applied over multiple domains:
  - Medical (golden standard for trauma assessment [43])
  - Archaeology [28]
  - Paleontology [23,44]
  - Material Science [50]

- Synchrotron facilities allow achieving **higher spatial and time resolutions** if compared to conventional X-Ray sources [16]

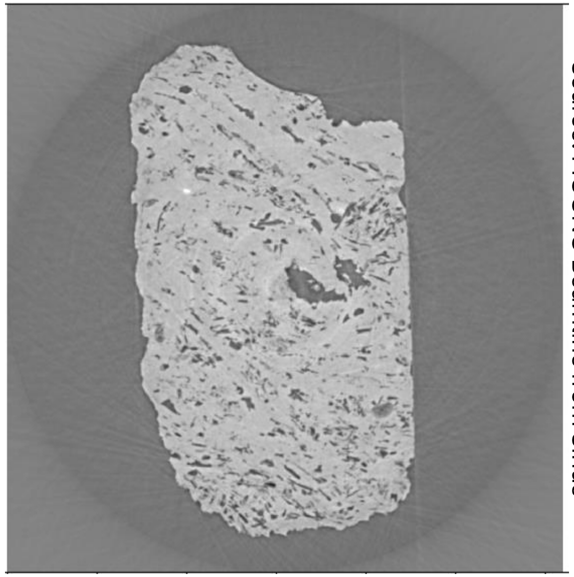

Source: Hamm et al. [23]

Tyrannosaurus rex left dentary

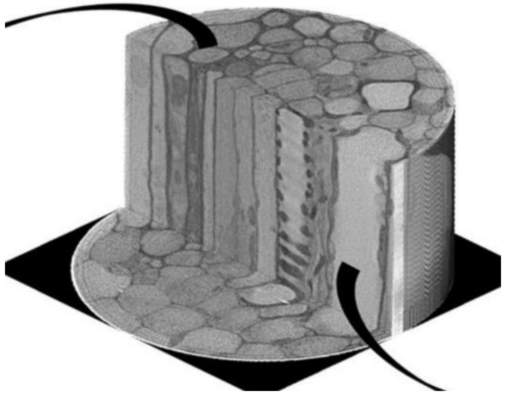


Source: Karjalainen et al. [28]

14th Century fabric


Source: MOGNO Beamline from Sirius

Carbonatic Rock


Source: Polo C. C. et al.*

Plant Tissue


Source: Wikipedia

Human body (Axial plane)

- High doses of radiation may be harmful for health [9,40] while for radiation-sensitive samples it may cause damages to it and directly impact the experimental results [36]

- Development of Low-dose Computed Tomography (LDCT) techniques is crucial
  - As Low As Reasonably Achievable (ALARA) principle

- Lower dose = lower CT image quality [17,49,51]
  - Higher noise
  - Lower contrast

- Methods to enhance LDCT image quality is crucial

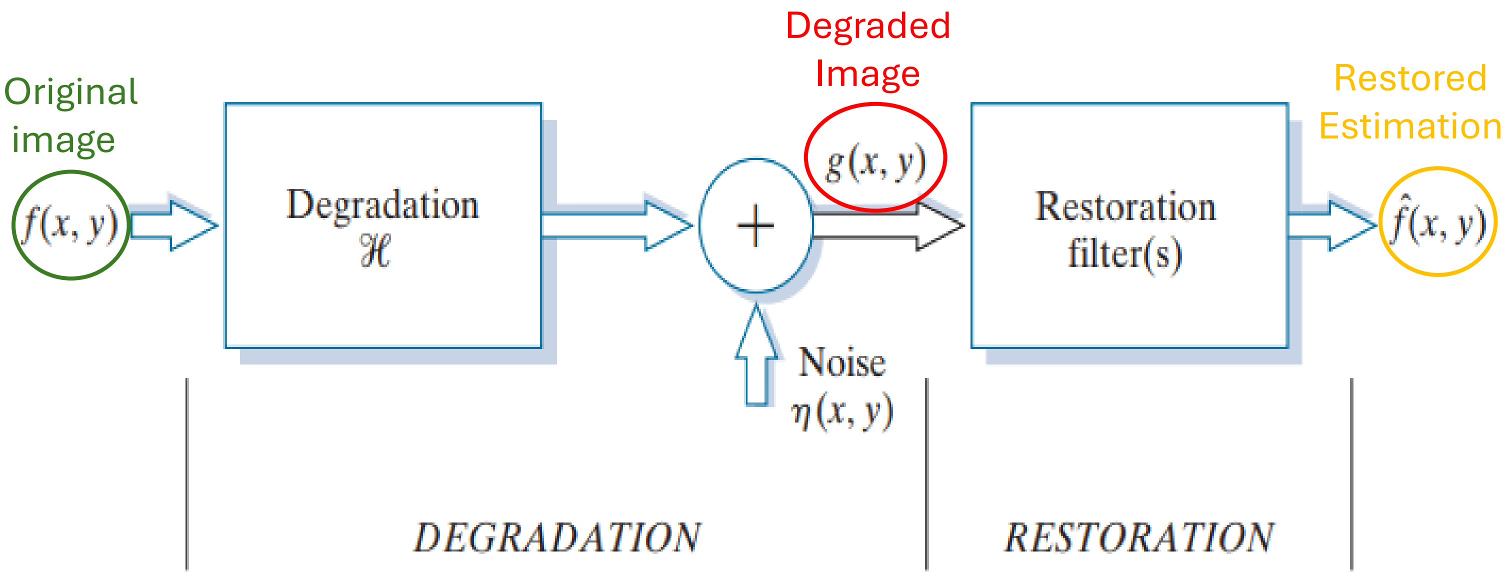
Low-dose

Normal-dose

Source: Stewart Bushong**

- The problem can be defined as a Degradation/Restoration problem

- **Objective**: Developed a way to restore the image as similar as possible to the original



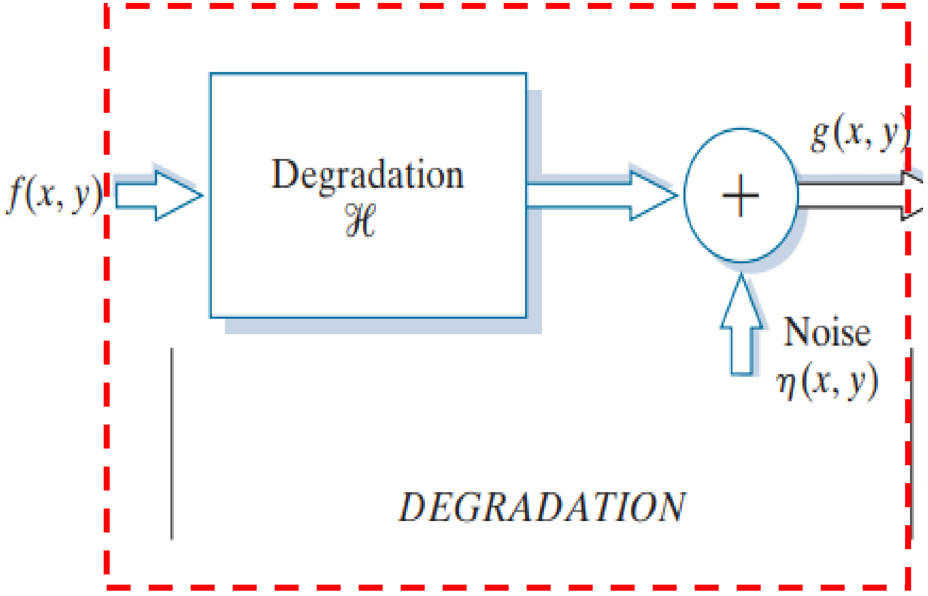
Source: Gonzales R. C. and Woods R. E. [20]

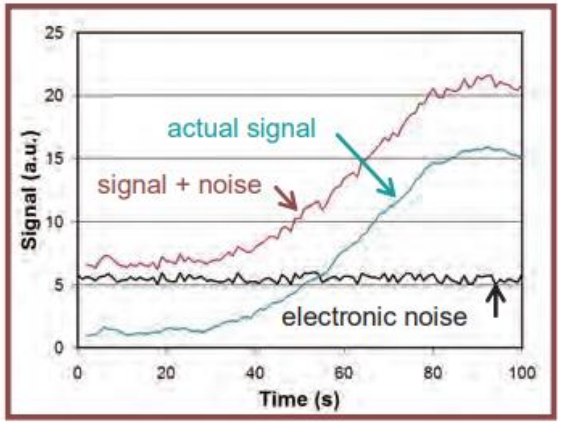
- The problem can be defined as a Degradation/Restoration problem

- **Objective**: Developed a way to restore the image as similar as possible to the original

- CT images are susceptible to various sources of noise [9]:
  - Grain noise
  - Quantum noise
  - Anatomical noise
  - Structure noise
  - Electronic noise

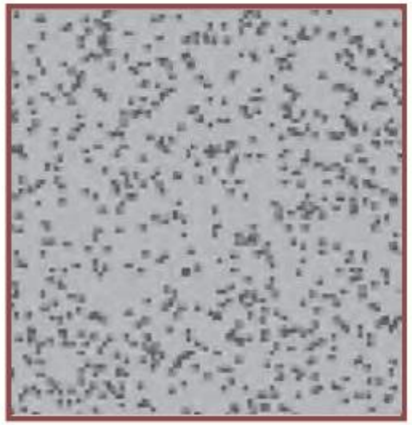
- Low-dose exposure worsens image quality





Electronic Noise

Source: Bushberg et al. [9]



Grain Noise

Source: Bushberg et al. [9]

Source: Bushberg et al. [9]

Structured Noise

Source: Bushberg et al. [9]

Anatomical Noise

Quantum mottle

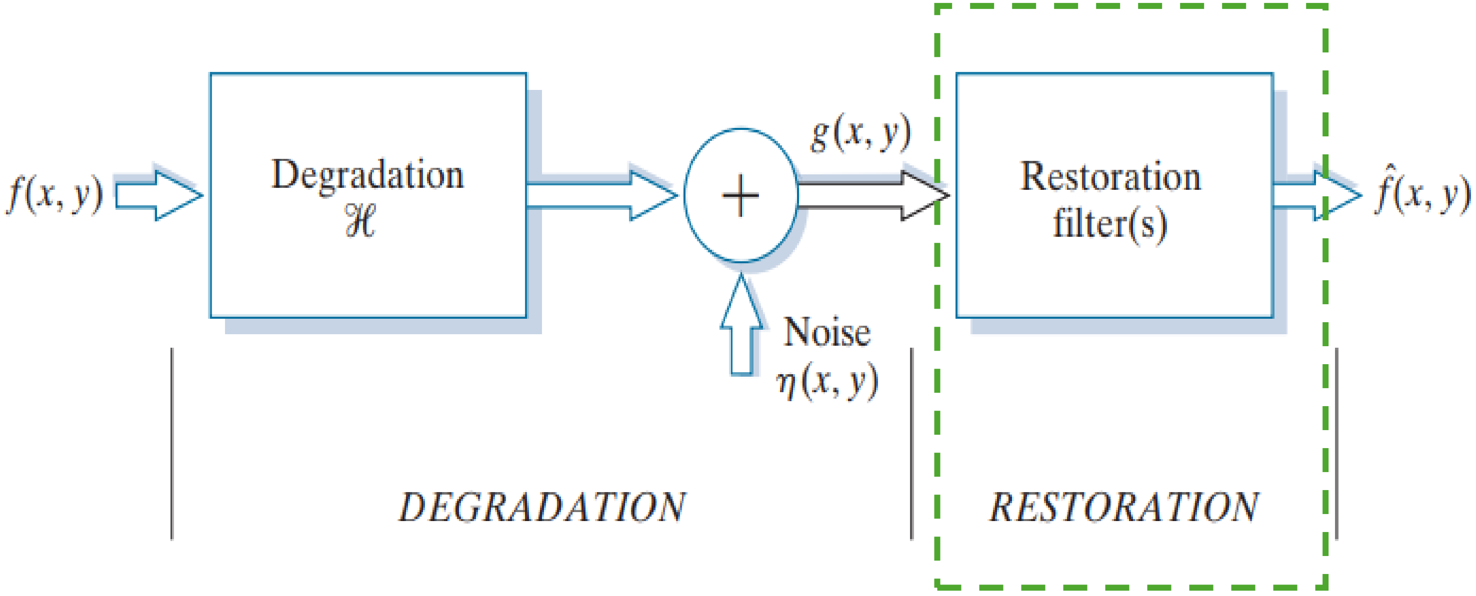
Optimal image

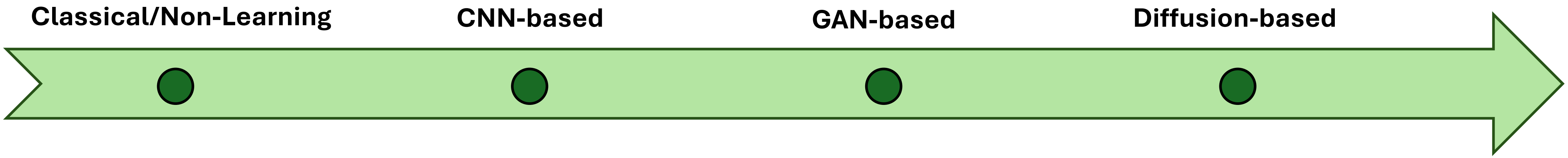
Source: Stewart Bushong*

- The problem can be defined as a Degradation/Restoration problem

- **Objective**: Developed a way to restore the image as similar as possible to the original

- Solutions are typically classified as [36, 42]:
  - Raw data filtering
  - Iterative reconstruction
  - **Post-processing algorithms**

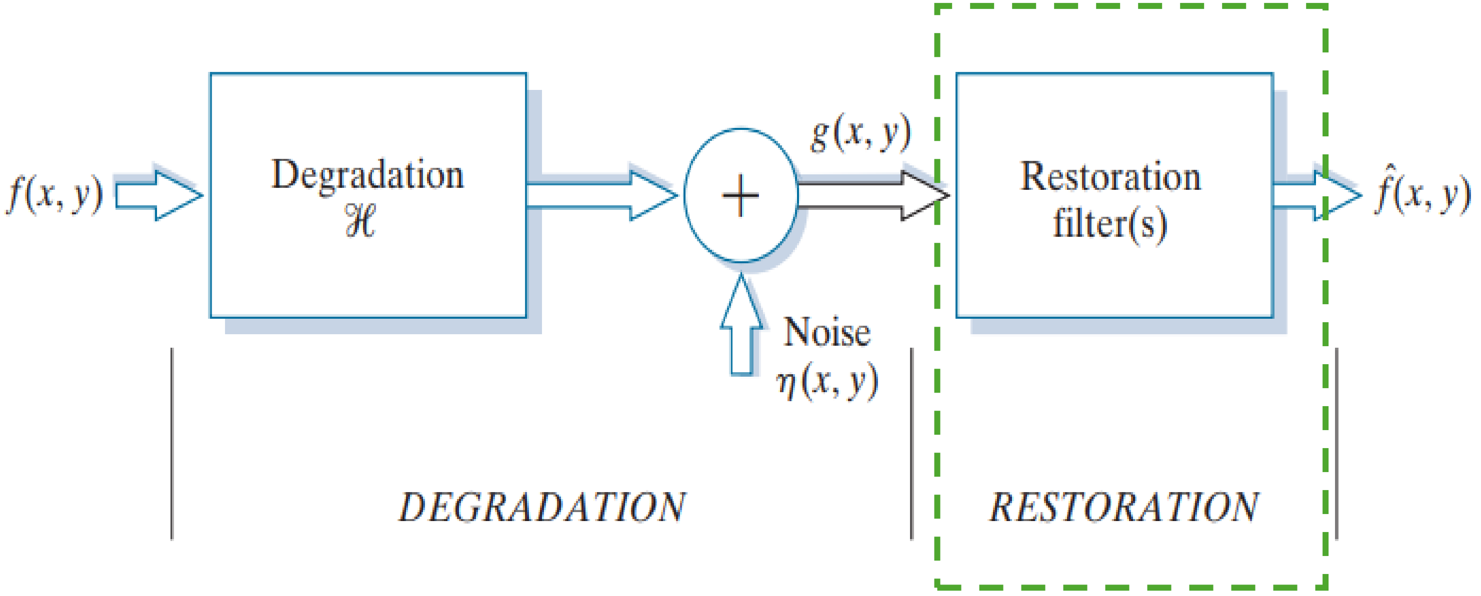




- The problem can be defined as a Degradation/Restoration problem

- **Objective**: Developed a way to restore the image as similar as possible to the original

- Solutions are typically classified as [36, 42]:
  - Raw data filtering
  - Iterative reconstruction
  - **Post-processing algorithms**



Source: Gonzales R. C. and Woods R. E. [20]

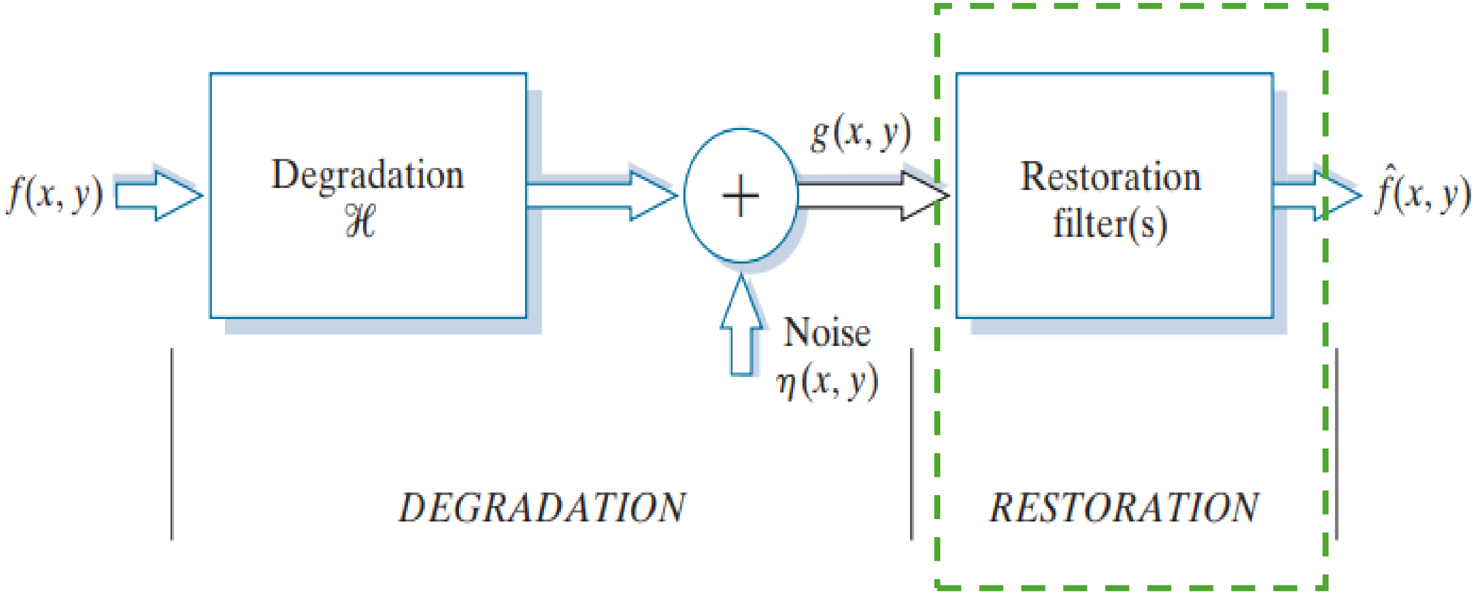
Classical/Non-Learning — **CNN-based** — **GAN-based** — **Diffusion-based**

1. Performance requires improvement [17, 56]

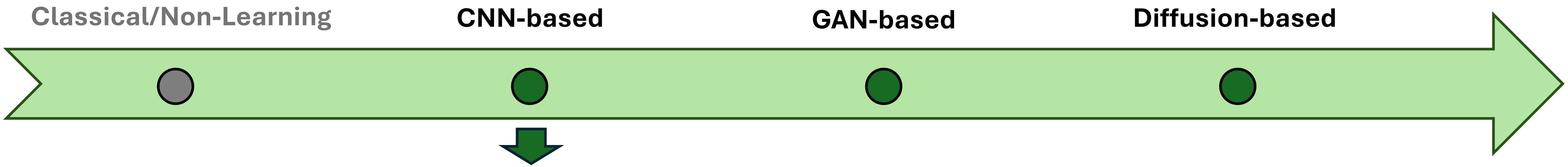
- The problem can be defined as a Degradation/Restoration problem

- **Objective**: Developed a way to restore the image as similar as possible to the original

- Solutions are typically classified as [36, 42]:
  - Raw data filtering
  - Iterative reconstruction
  - **Post-processing algorithms**



Source: Gonzales R. C. and Woods R. E. [20]

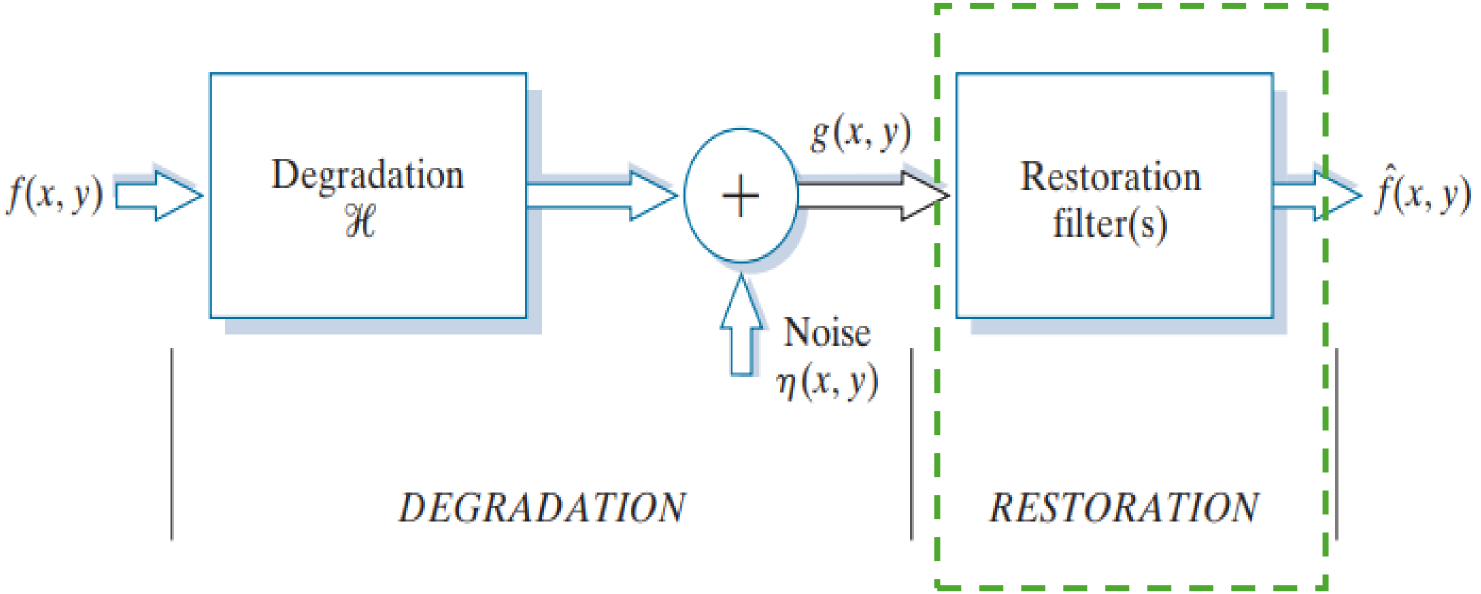


1. Performance requires improvement [17, 56]

1. Good Performance
2. Causes over-smoothing
3. Poor generalization for unseen dose levels

[17 ,18]

- The problem can be defined as a Degradation/Restoration problem
- **Objective**: Developed a way to restore the image as similar as possible to the original
- Solutions are typically classified as [36, 42]:
  - Raw data filtering
  - Iterative reconstruction
  - **Post-processing algorithms**



Source: Gonzales R. C. and Woods R. E. [20]

Classical/Non-Learning

1. Performance requires improvement [17, 56]

CNN-based

1. Good Performance
2. Causes over-smoothing
3. Poor generalization for unseen dose levels [17 ,18]

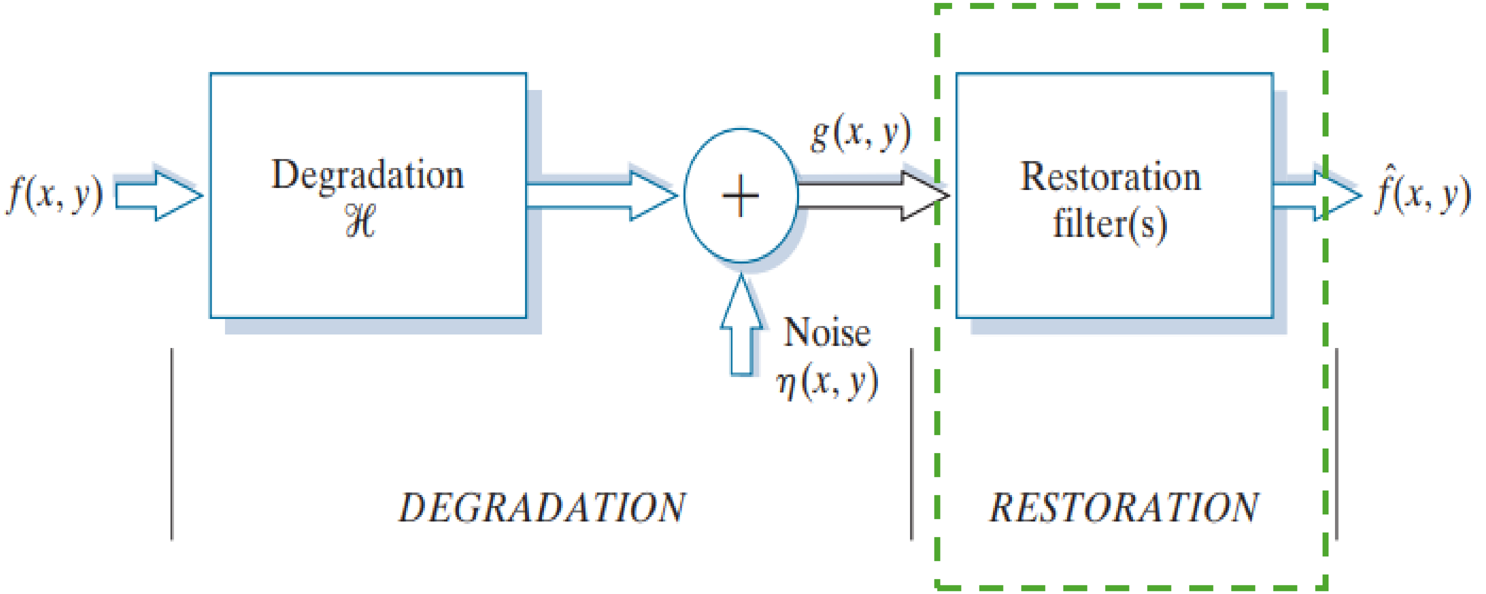
GAN-based

1. Better detail preservation
2. Better generalization
3. Unstable train
4. Value shift

[17, 18]

Diffusion-based

- The problem can be defined as a Degradation/Restoration problem
- **Objective**: Developed a way to restore the image as similar as possible to the original
- Solutions are typically classified as [36, 42]:
  - Raw data filtering
  - Iterative reconstruction
  - **Post-processing algorithms**



Source: Gonzales R. C. and Woods R. E. [20]

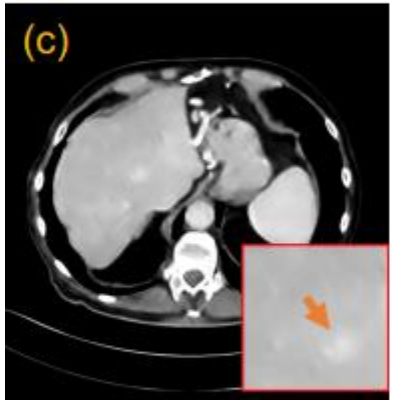
Classical/Non-Learning | **CNN-based** | **GAN-based** | **Diffusion-based**

Classical/Non-Learning:
1. Performance requires improvement [17, 56]

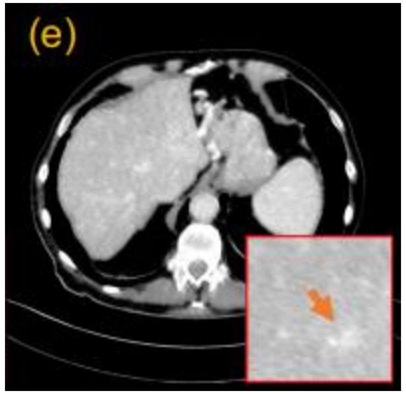
CNN-based:
1. Good Performance
2. Causes over-smoothing
3. Poor generalization for unseen dose levels [17 ,18]

GAN-based:
1. Better detail preservation
2. Better generalization
3. Unstable train
4. Value shift [17, 18]

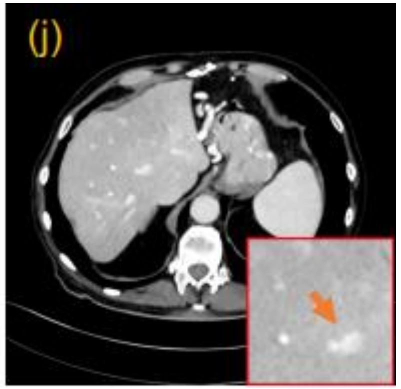
Diffusion-based:
1. Promising results for denoising in other domains when compared to CNN and GANS
2. Already explored for LDCT [17,18,35, 56]
3. Not extensively explored yet
4. Never applied to synchrotron LDCT denoising



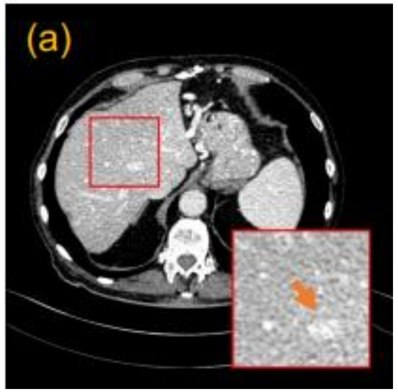
RED-CNN



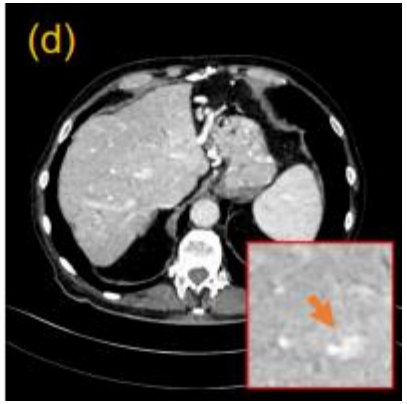
WGAN-VGG



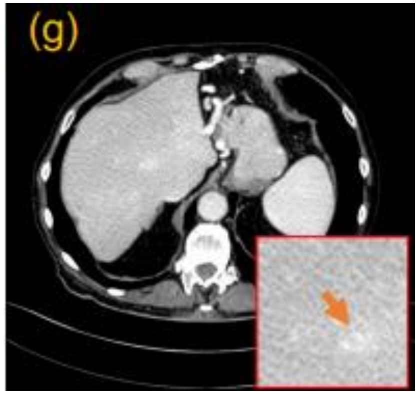
CoreDiff-10



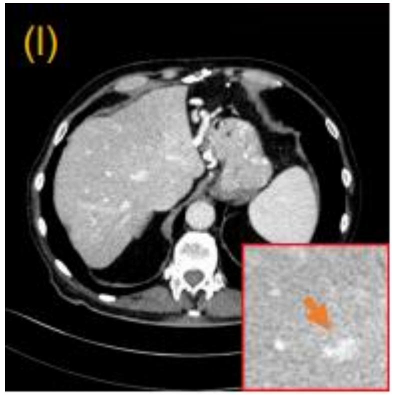
NDCT



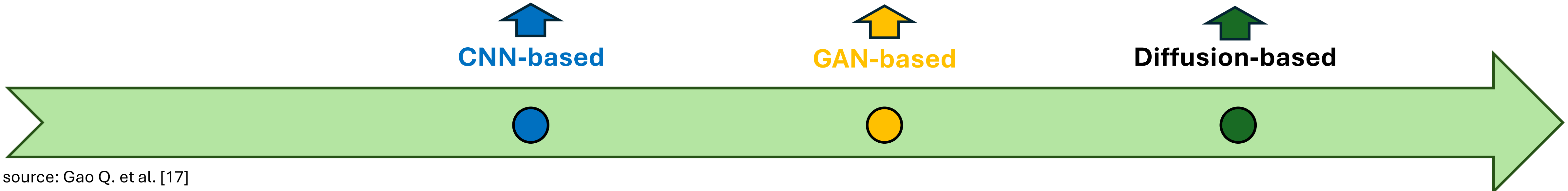
PDF-RED-CNN



DU-GAN



CoreDiff+OSLu -10



*Images source: Gao Q. et al. [17]

Can diffusion models perform denoising tasks on LDCT reconstructed data toward increasing the quality of synchrotron LDCT?

Can diffusion models trained over CT medical images be directly repurposed to perform denoising tasks on LDCT synchrotron images?

Does finetunning a model trained over CT medical images with synchrotron CT images enhances the acquired results?

Evaluate diffusion-based generative models for denoising synchrotron LDCT

Assess the generalization of diffusion models trained on medical LDCT images for denoising synchrotron LDCT images

Contribute with a methodology based on diffusion generative models for enhancing synchrotron LDCT image quality

Explore taking advantage of medical LDCT datasets to train models to be used over synchrotron LDCT images

Eletromagnetic Radiation (EM)

- Radiation is “energy that travels through space or matter” [9]

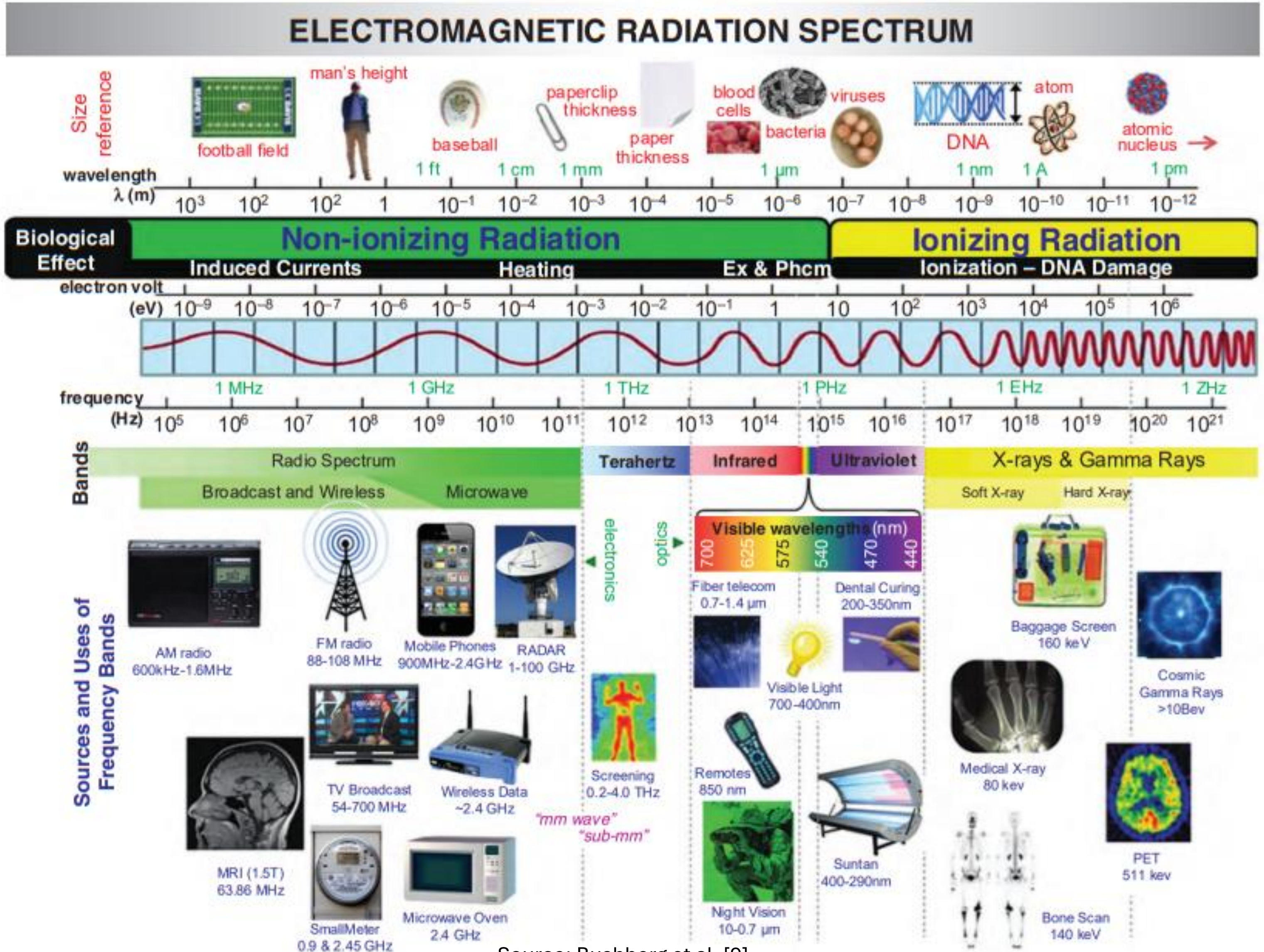
- Due to the “wave-particle duality” from quantum mechanics EM can be described as both waves and particles called *photon*

- Characterized by:
  - Wavelength (λ)
  - Frequency (v)
  - Energy (E)

$$E = h\nu = \frac{hc}{\lambda}$$

- Divided in groups according to those characteristics

- Ionization may occur when *photon* interacts with molecules/atom depending on:
  - *Photon* energy
  - Target molecule/atom



Source: Bushberg et al. [9]

Eletromagnetic Radiation (EM)

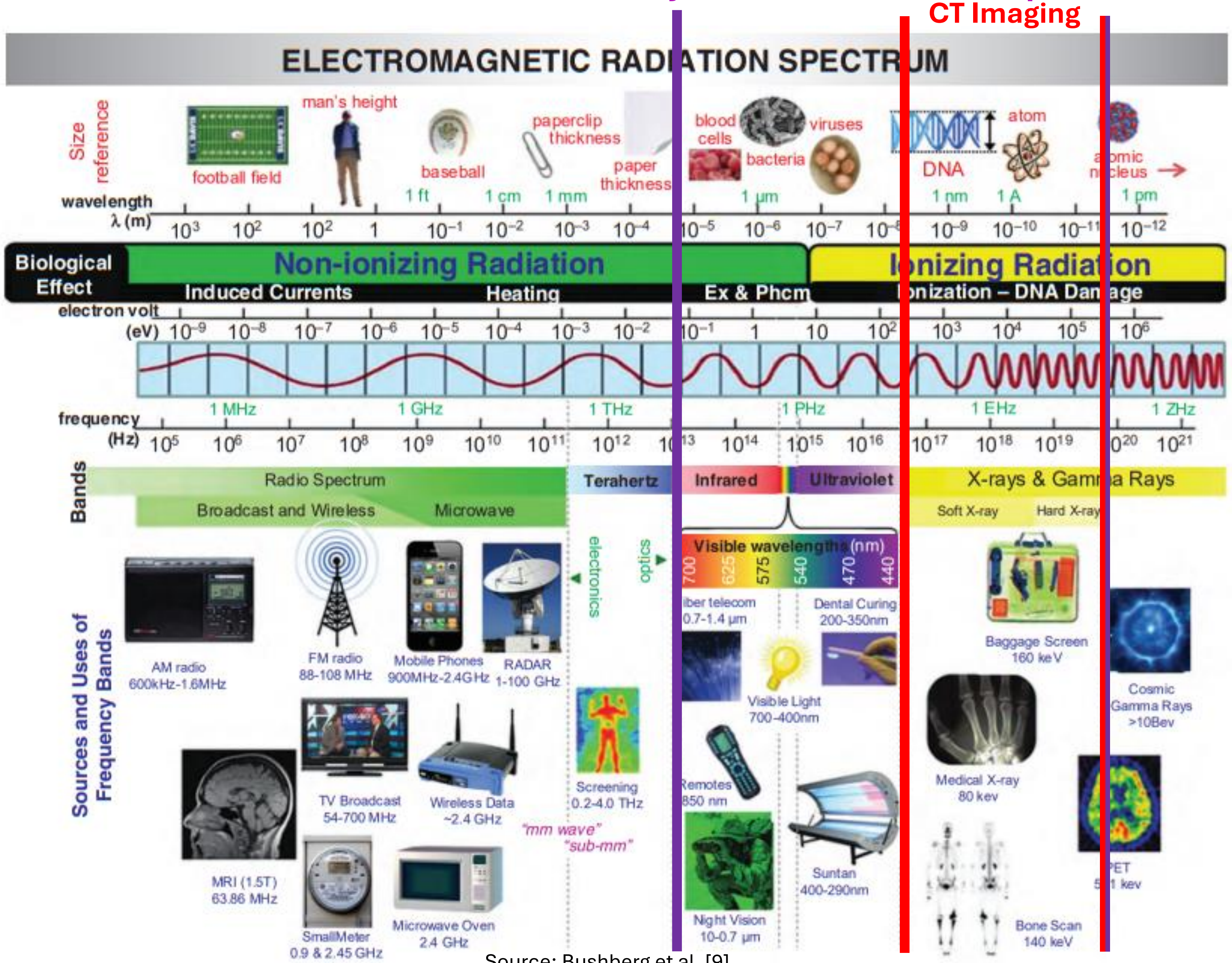
- Radiation is "energy that travels through space or matter" [9]
- Due to the "wave-particle duality" from quantum mechanics EM can be described as both waves and particles called *photon*
- Characterized by:
  - Wavelength (λ)
  - Frequency (ν)
  - Energy (E)

$$E = h\nu = \frac{hc}{\lambda}$$

- Divided in groups according to those characteristics
- Ionization may occur when *photon* interacts with molecules/atom depending on:
  - *Photon* energy
  - Target molecule/atom

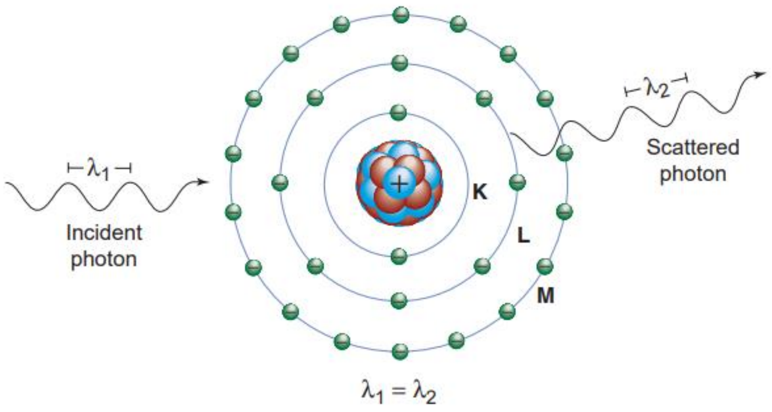
Synchrotron Radiation Spectrum

CT Imaging

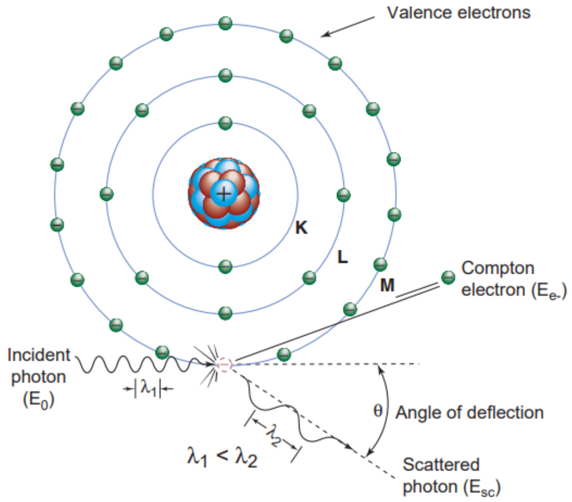


Source: Bushberg et al. [9]

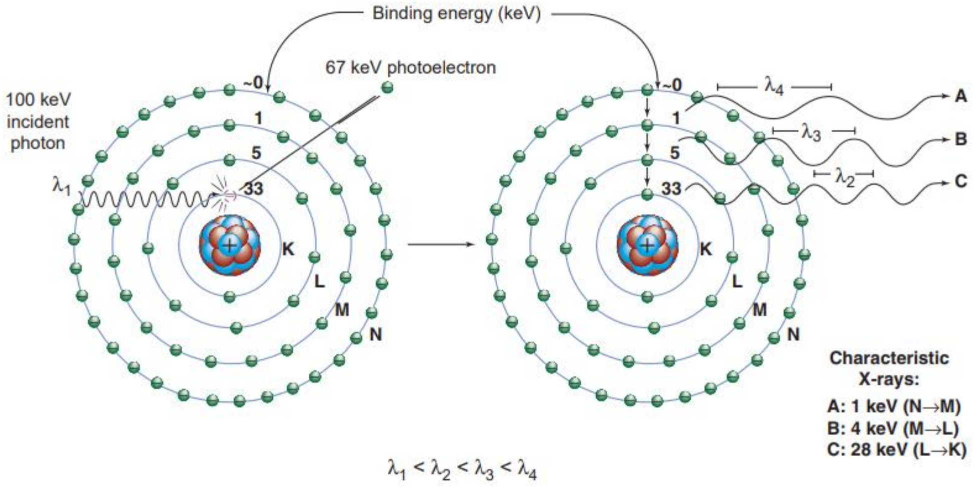
- Types of interaction between EM radiation and matter [9]:
  - Rayleigh Scattering
  - Compton Scattering
  - Photoelectric Absorption
  - Pair Production (Only on high-energies)



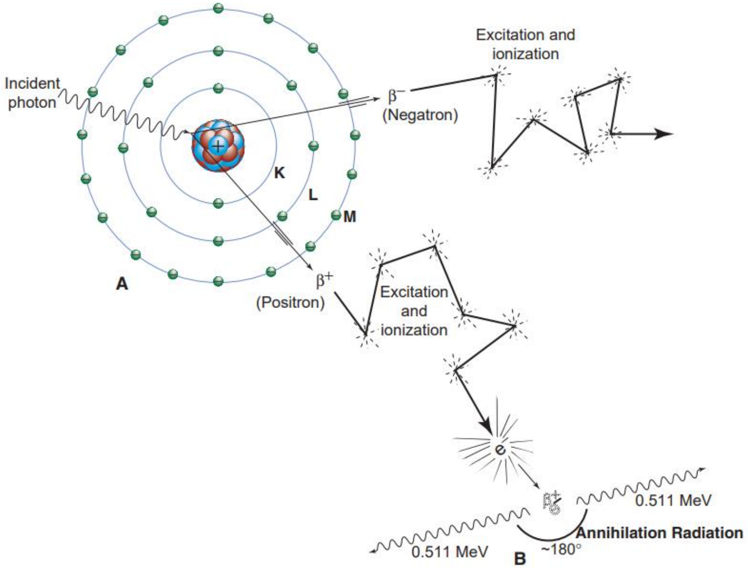
Rayleight Scattering



Compton Scattering



Photoelectric Absorption



Pair Production

Source: Bushberg et al. [9]

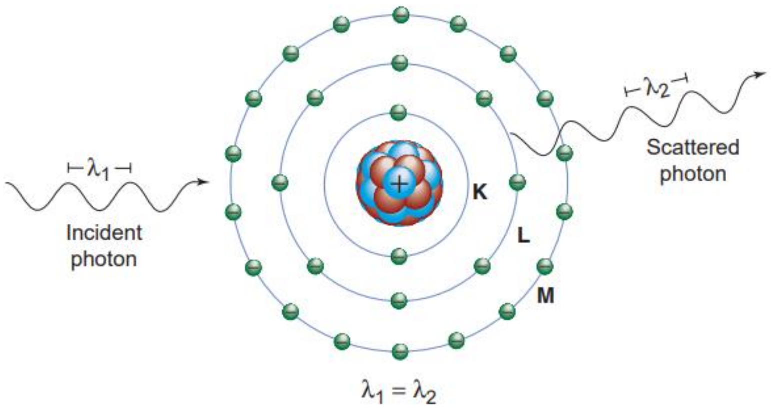
- Types of interaction between EM radiation and matter [9]:
  - Rayleigh Scattering
  - Compton Scattering
  - Photoelectric Absorption
  - Pair Production (Only on high-energies)

Combined occurrence

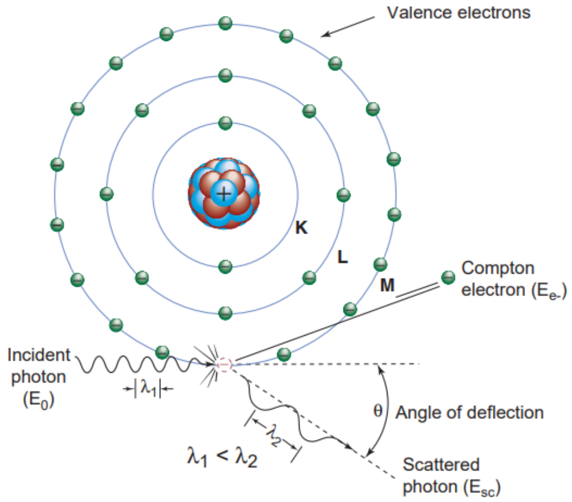
EM Attenuation

Varies with material composition and thickness

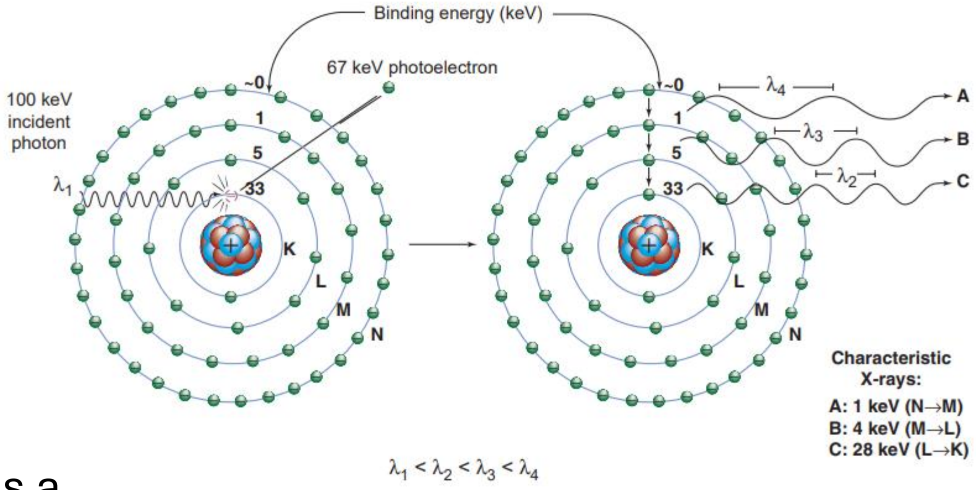
EM Attenuation can be defined as a **Proportion of atoms per volume [9]**



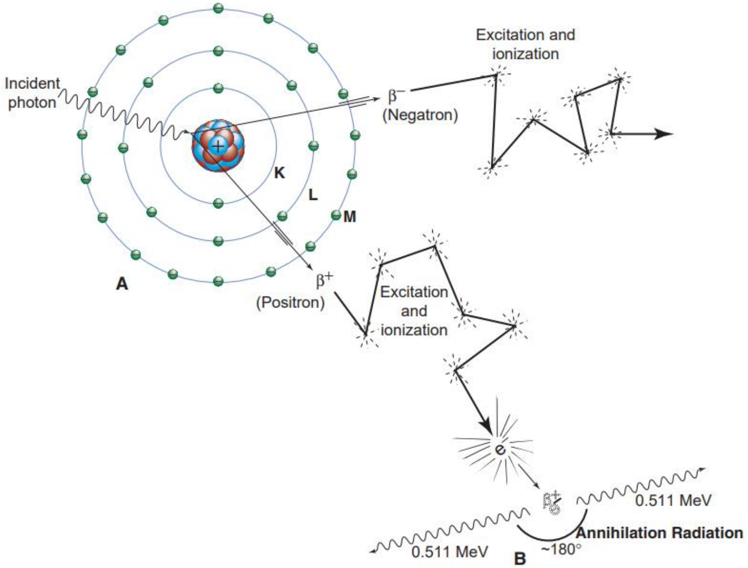
Rayleight Scattering



Compton Scattering



Photoelectric Absorption



Pair Production

Source: Bushberg et al. [9]

Radiation Interaction with Matter

- Types of interaction between EM radiation and matter [9]:
  - Rayleigh Scattering
  - Compton Scattering
  - Photoelectric Absorption
  - Pair Production (Only on high-energies)

Combined occurrence

EM Attenuation

Varies with material composition and thickness

EM Attenuation can be defined as a **Proportion of atoms per volume [9]**

$$\mu = \mu_{\text{Rayleigh scatter}} + \mu_{\text{photoelectric effect}} + \mu_{\text{Compton scatter}} + \mu_{\text{pair production}}$$

$$\text{Mass Attenuation Coefficient} = \frac{\mu}{\rho}$$

Material density

- Types of interaction between EM radiation and matter [9]:
  - Rayleigh Scattering
  - Compton Scattering
  - Photoelectric Absorption
  - Pair Production (Only on high-energies)

Combined occurrence

EM Attenuation

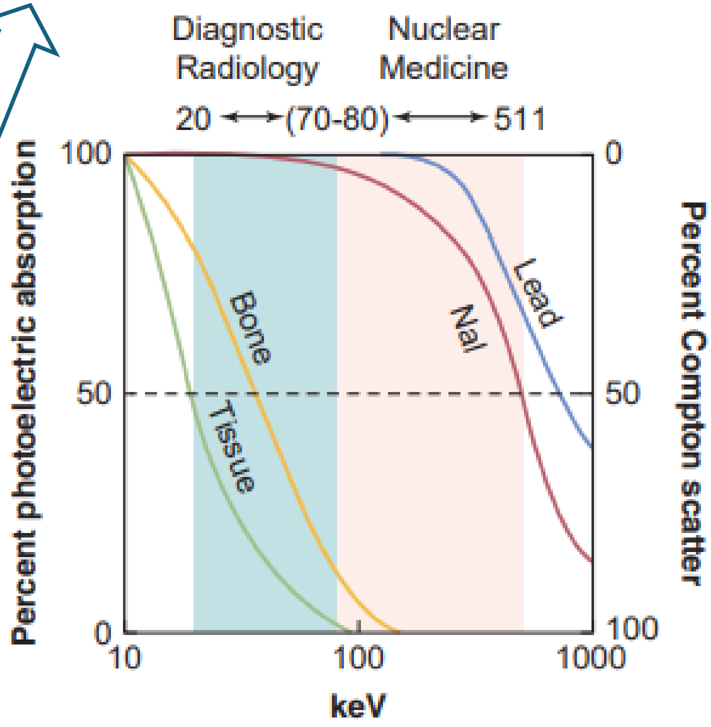
Varies with material composition and thickness

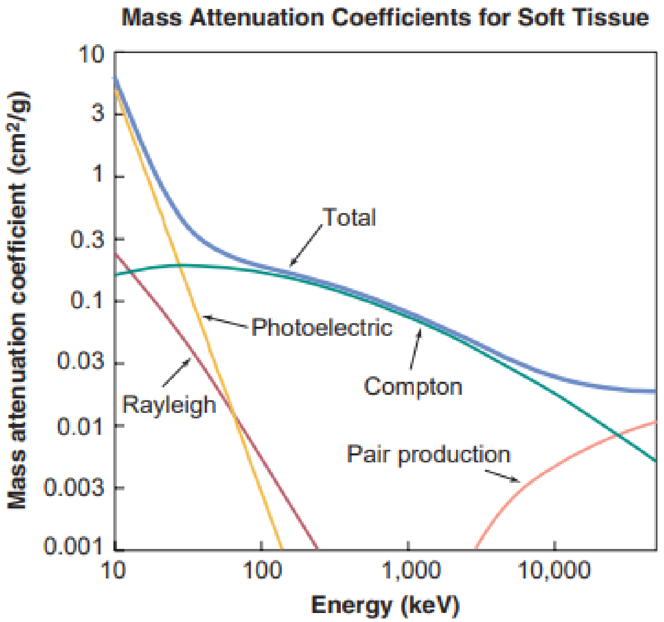
EM Attenuation can be defined as a **Proportion of atoms per volume [9]**

$$\mu = \mu_{\text{Rayleigh scatter}} + \mu_{\text{photoelectric effect}} + \mu_{\text{Compton scatter}} + \mu_{\text{pair production}}$$

$$\text{Mass Attenuation Coefficient} = \frac{\mu}{\rho}$$

Material density





Source: Bushberg et al. [9]

Radiation Dose

- Radiation dose is the energy deposited in the irradiated material [9]
- Fluence (ϕ) and flux (ψ)
  - Synchrotron facilities have a higher flux than medical X-Ray sources

$$\phi = \frac{Photons}{Area} \qquad \psi = \phi \times E$$

Very simplified way to calculate, as it don't consider:

- Poly-energetic behavior of the radiation source
- Material heterogeneity and morphology
- Radiation-matter interaction variation
- Scattering causing re-interaction
- Surroundings

High complexity on precisely calculating dose

Due to that complexity, dose calculation varies depending on application

- Absorbed Dose (D)

$$D = \frac{E}{m}$$

Irradiated material mass

- Radiation protection and Medical Imaging application commonly uses:
  - Equivalent Dose (H)
    - Adds a weighting factor based on the radiation type
  - Effective Dose (E_{dose})
    - Adds a weighting factor based on each material being irradiated

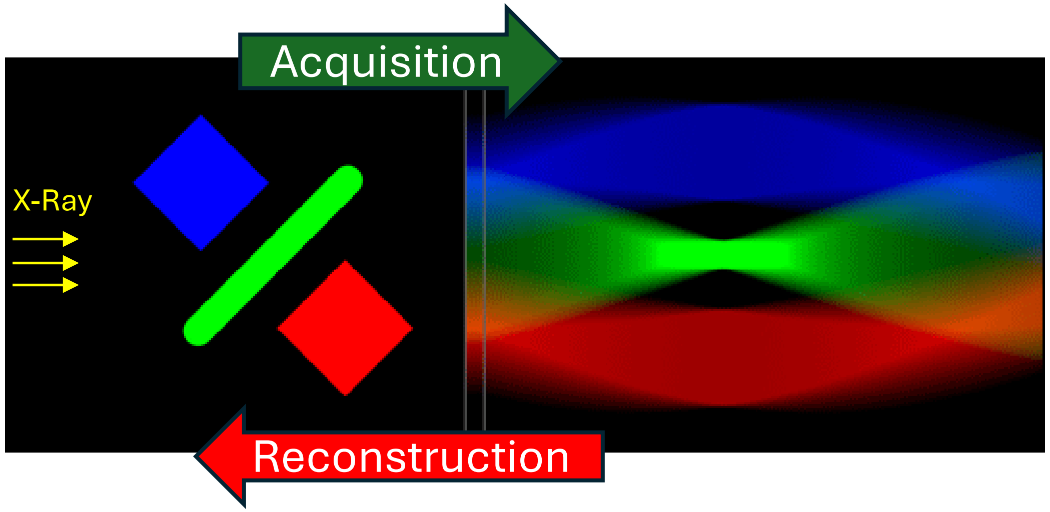
$$H = D \times w_R \qquad E_{dose} = \sum_T (w_T \times H_T)$$

Radiation effectiveness on causing biological damage

Material sensitivity

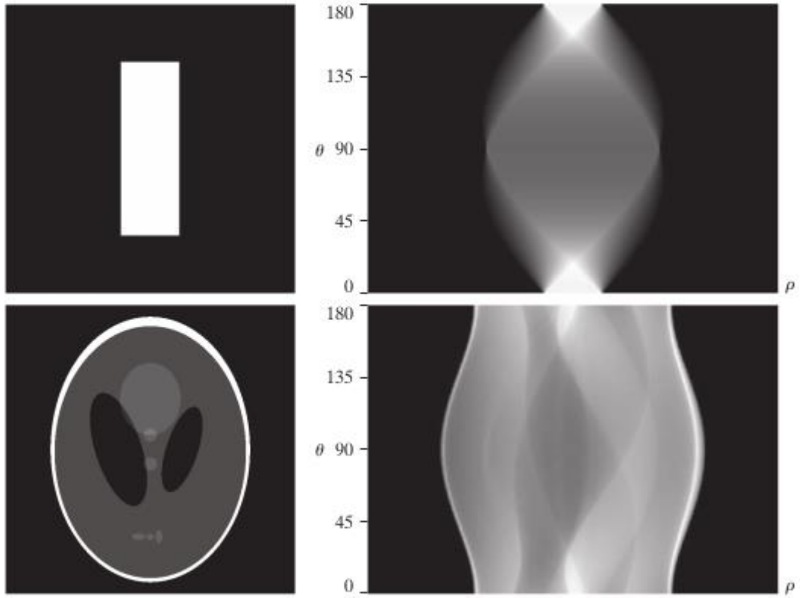
X-Ray Computed Tomography (CT)

- X-Ray Computer Tomography can be summarized in two stages acquisition and reconstruction [31 , 36]
- **Acquisition**: Capture 2D transmission projection images of an object from various angles around a common axis
- **Reconstruction**: Apply a computational reconstruction method to restore object's 3D morphology

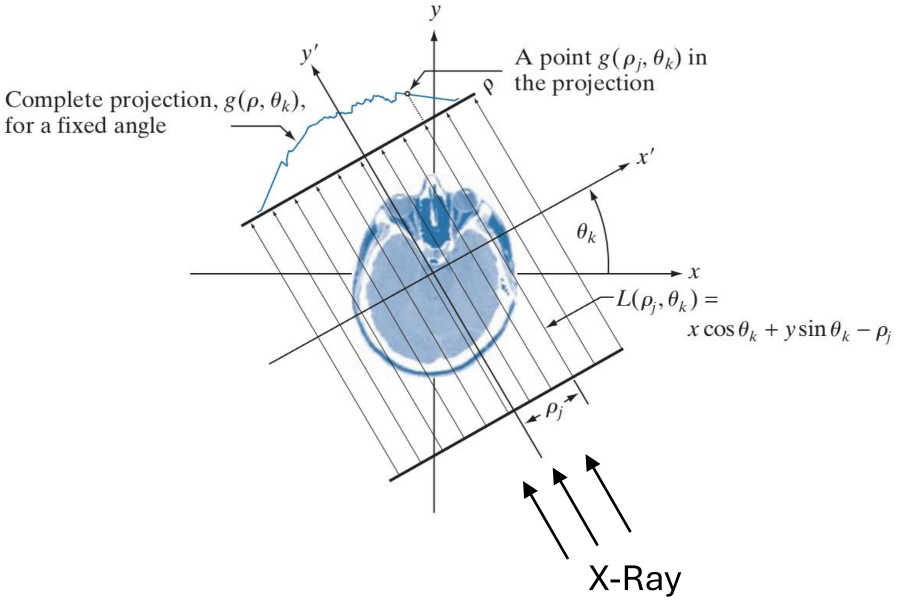


Source: Wikipedia

Credits: MOGNO beamline staff



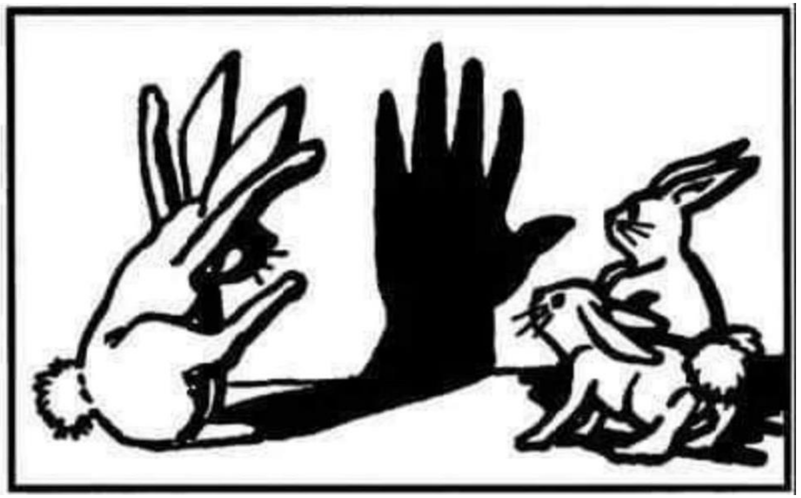
Source: Gonzales R. C. and Woods R. E. [20]



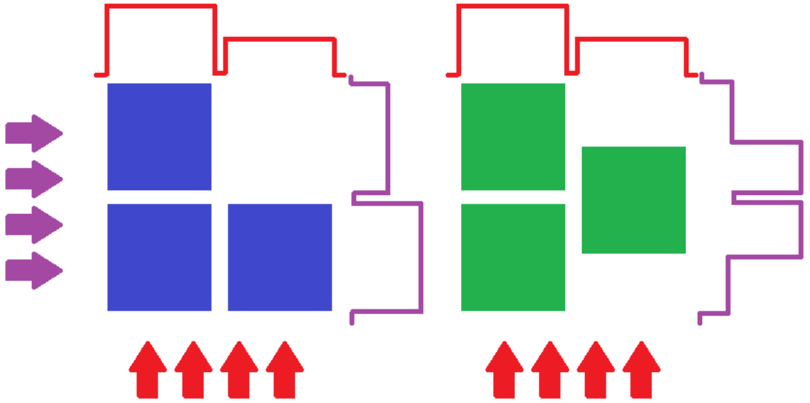
Source: Gonzales R. C. and Woods R. E. [20]

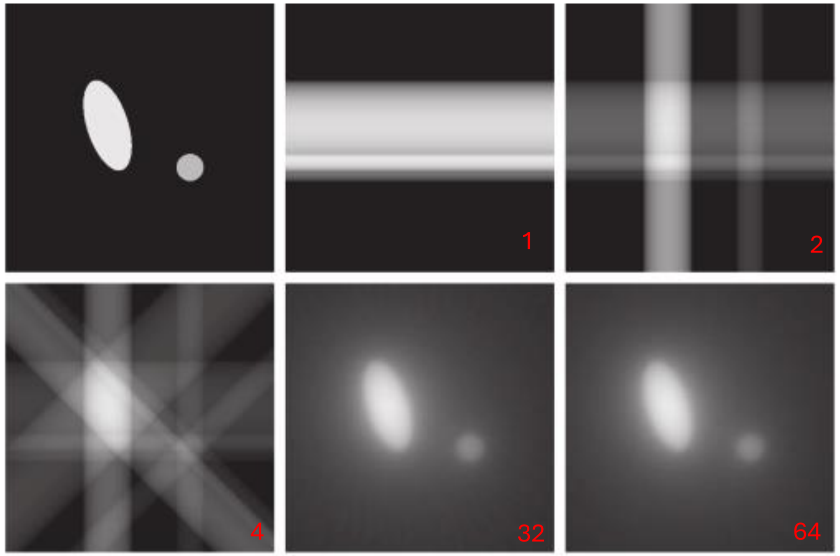
X-Ray Computed Tomography (CT)

- X-Ray Computer Tomography can be summarized in two stages acquisition and reconstruction [31 , 36]
- **Acquisition**: Capture 2D transmission projection images of an object from various angles around a common axis
- **Reconstruction**: Apply a computational reconstruction method to restore object's 3D morphology
  - Inverse problem
  - Different object may cause same projection
- The more acquired projections the better the reconstruction
  - In theory, with ∞ projections, it would be possible to invert the Radon transform exactly



Source: "La Découverte de l'ombre" (Roberto Casati)

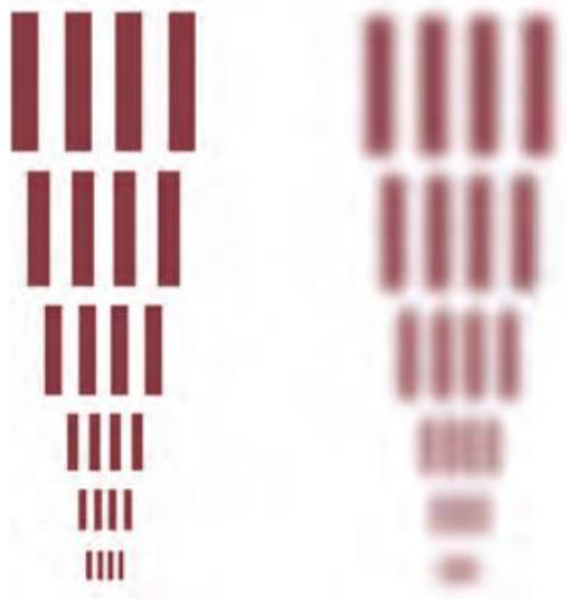




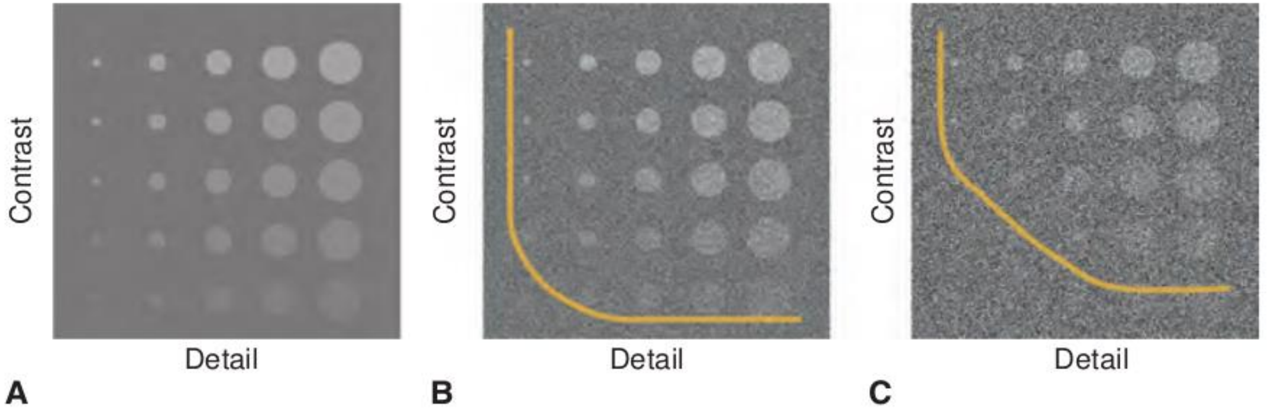
Source: Gonzales R. C. and Woods R. E. [20]

- According to Bushberg et al. [9], CT image quality is strongly bounded to:
    - Spatial Resolution
    - Contrast Resolution
    - Temporal Resolution

- **Spatial Resolution**:
    - Ability to distinguish two objects of different densities
    - Determines edge sharpness and detail clarity
    - Related to how much of real space is represented by a pixel/voxel
- **Contrast Resolution**:
    - Ability to differentiate objects with similar densities using grayscale values
    - Emphasizes distinction between similarly shaded objects
- **Temporal Resolution**:
    - How long CT image acquisition takes
    - Crucial for imaging moving objects

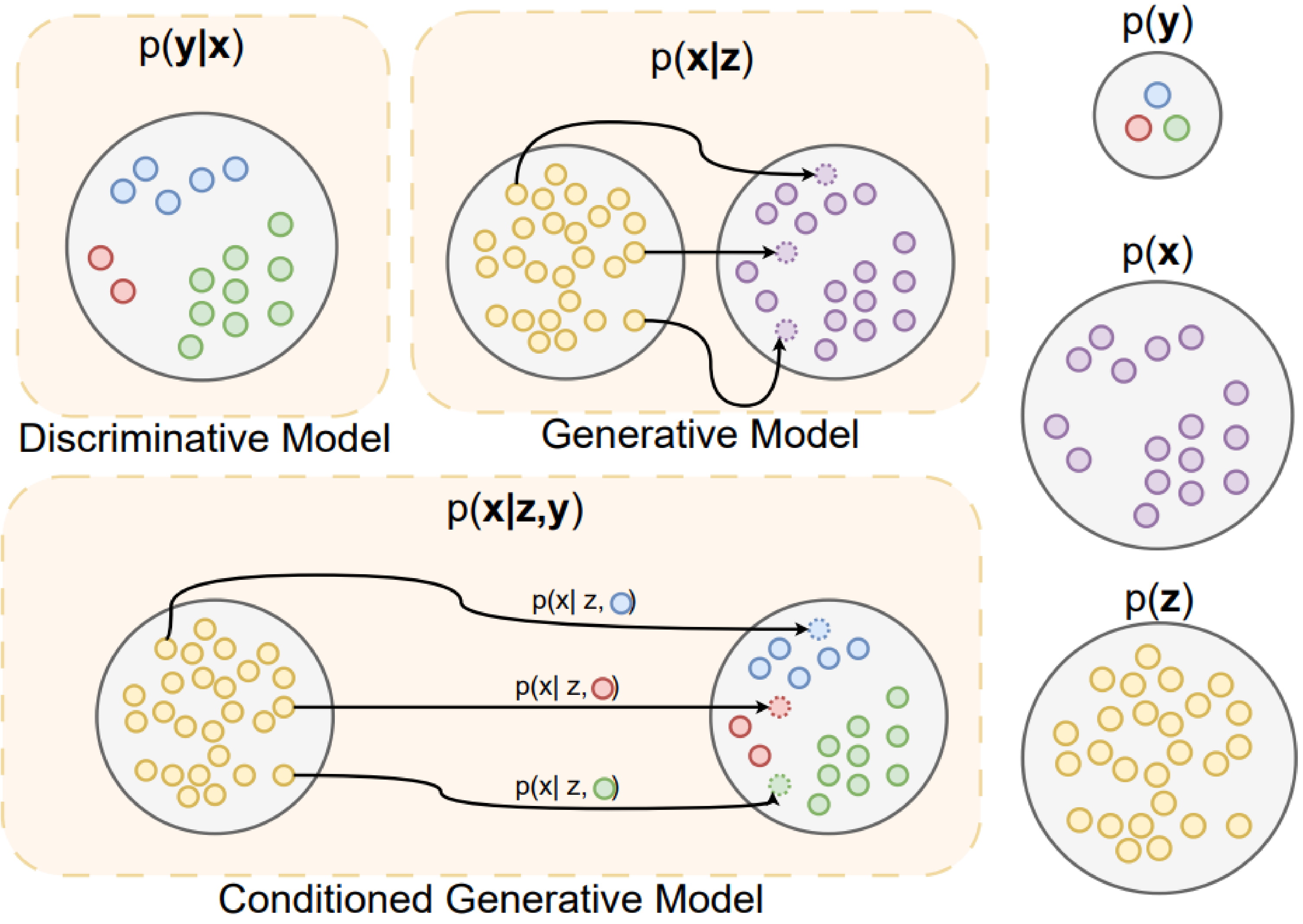


Source: Bushberg et al. [9]



Source: Bushberg et al. [9]

- **Discriminative Models**:
  - Learn a function capable of defining boundaries that distinguish which class a sample fits
- **Generative Models** :
  - Learn how to transform a latent space variable (**z**) into a data space variable (**x**)
- **Conditioned Generative Models** :
  - The same as generative models but condition the transformation to guide the transformation



Source: Author creation

- **Discriminative Models**:
  - Learn a function capable of defining boundaries that distinguish which class a sample fits
- **Generative Models** :
  - Learn how to transform a latent space variable (**z**) into a data space variable (**x**)
- **Conditioned Generative Models** :
  - The same as generative models but condition the transformation to guide the transformation

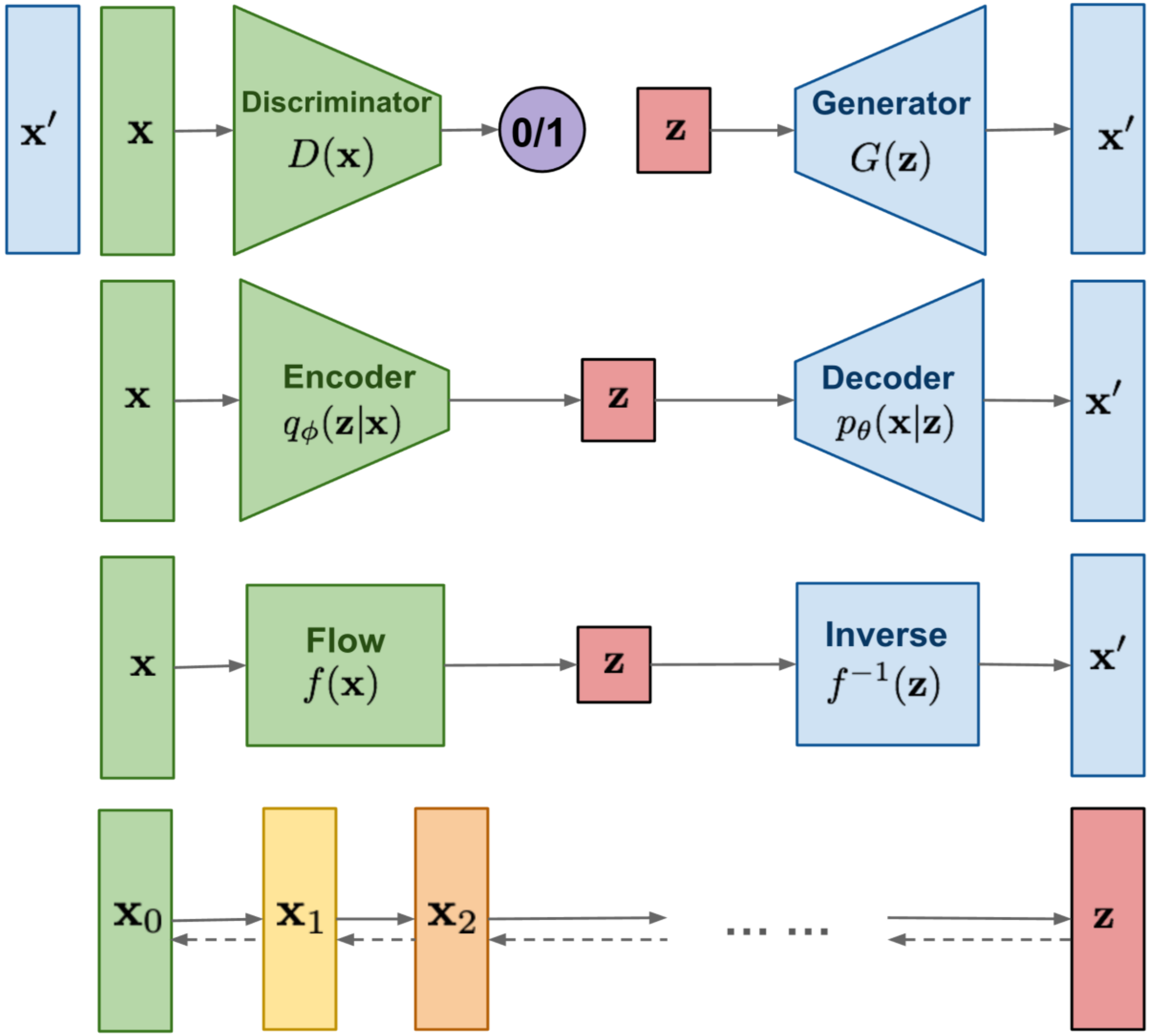
- Learning to map latent space variables (**z**) to data space variables (**x**) is the core idea of many generative models.

GAN: Adversarial training

VAE: maximize variational lower bound

Flow-based models: Invertible transform of distributions

Diffusion models: Gradually add Gaussian noise and then reverse



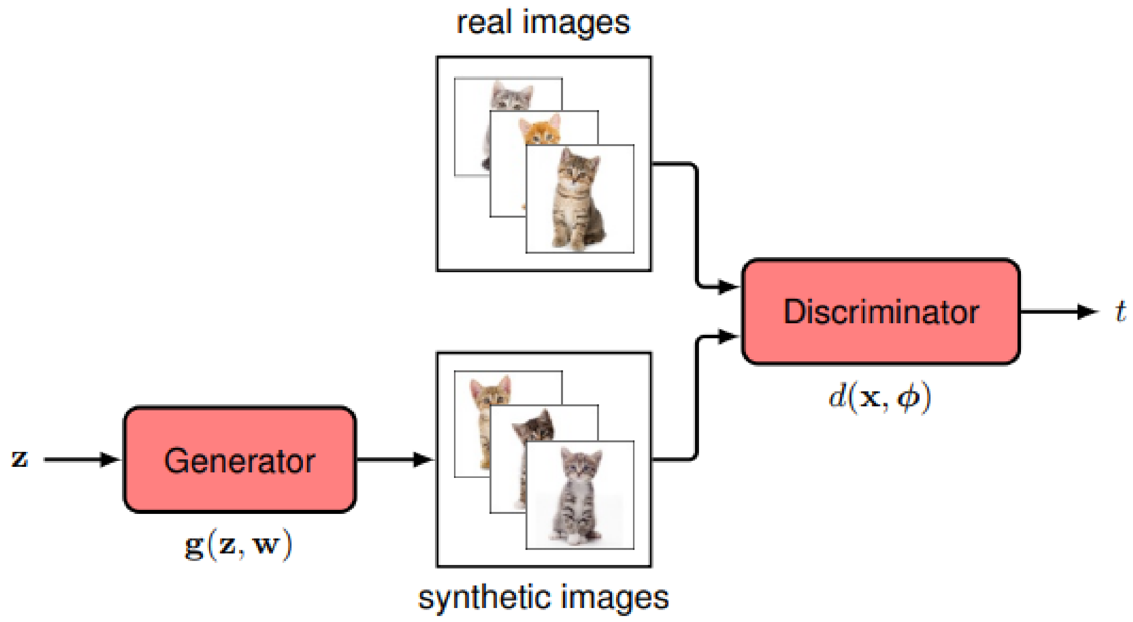
Source: Weng, Lilian [54]

- Generative Adversarial Networks (GANS) [21]
    - **Generator**
        - Generate realistic samples by transforming data from latent space p(**z**) to samples from data space p(**x**)
    - **Discriminator**
        - Distinguish fake and real samples

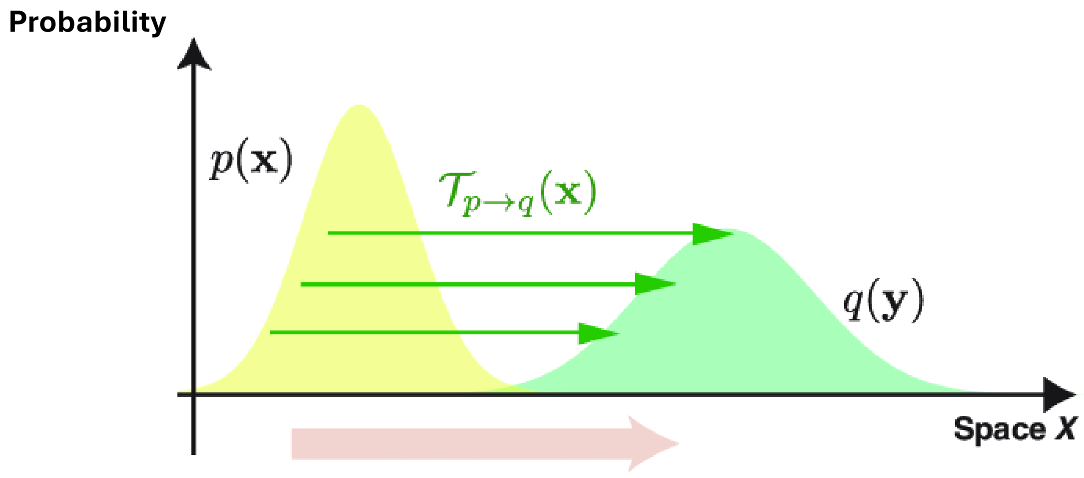
- Adversarial training:
    - Shared loss between models
    - Discriminator → Minimizes the error
    - Generator → Maximizes the error
    - Force both models to improve iteratively

- Training GANs can be challenging
    - Model collapse: Generator map p(**z**) to only a subgroup of p(**x**)
    - No clear progression metric

- Wasserstein Generative Adversarial Network (WGAN) [37]
    - More stability
    - Ensure the generator is Moving toward the desired distribution
    - Meaningful loss to show training progression



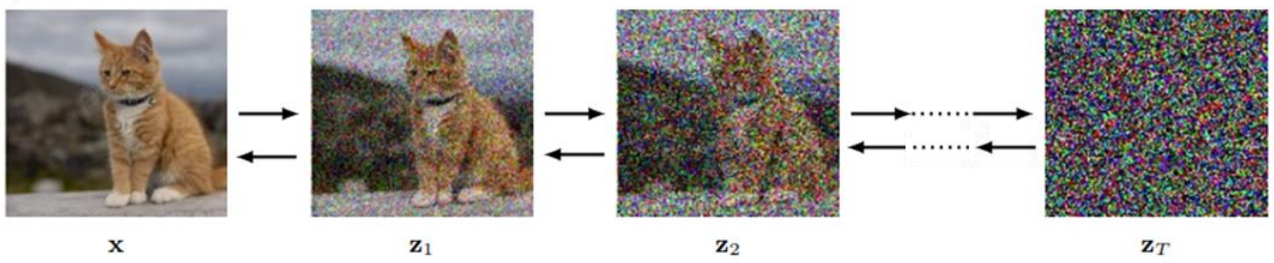
Source: Bishop and Bishop [5]

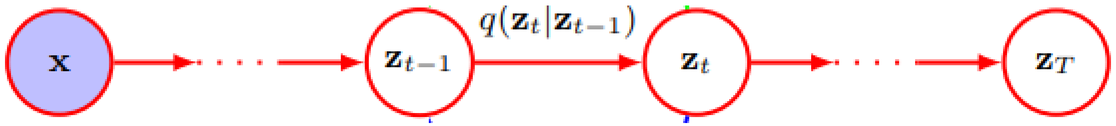


Source: Nakazato and Ito ***

Wasserstein distance

- Diffusion models [46] typically involve two stages:
  - **Forward encoder**
    - Gradually corrupts input data by adding noise to it
    - After several steps input data is transformed on a known noise-like distribution





Source: Bishop and Bishop [5]

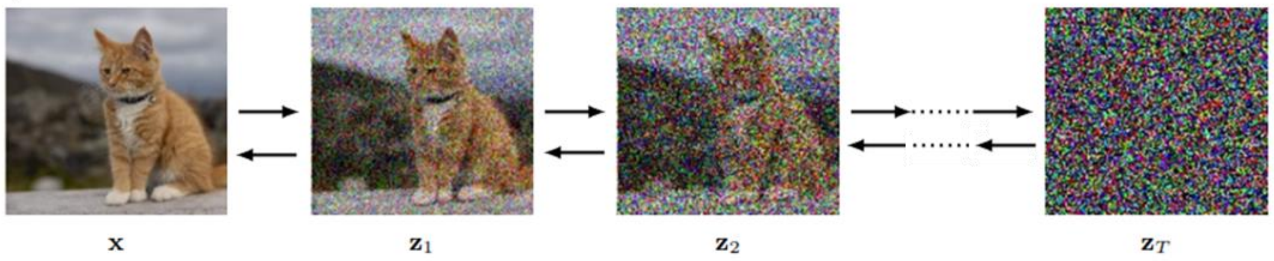
Noise add each step

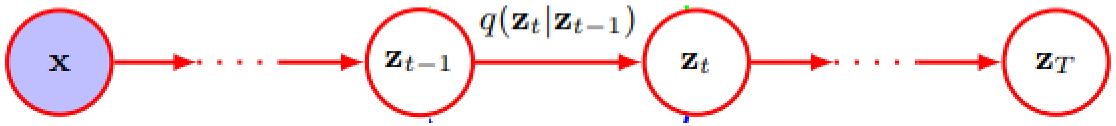
$$\boxed{\boldsymbol{\epsilon}_t \sim N(\boldsymbol{\epsilon}_t|0, \mathbf{I})}$$

$$\mathbf{z}_t = \sqrt{1-\beta_t}\mathbf{z}_{t-1} + \sqrt{\beta_t}\boldsymbol{\epsilon}_t$$

Noise distribution variance on step t

- Diffusion models [46] typically involve two stages:
  - **Forward encoder**
    - Gradually corrupts input data by adding noise to it
    - After several steps input data is transformed on a known noise-like distribution





Source: Bishop and Bishop [5]

$$\mathbf{z}_t = \boxed{\sqrt{1-\beta_t}}\,\mathbf{z}_{t-1} + \boxed{\sqrt{\beta_t}}\,\epsilon_t$$

Noise add each step

$$\epsilon_t \sim N(\epsilon_t|0, \mathbf{I})$$

Noise distribution variance on step t

$$\beta_1 < \beta_2 < \cdots < \beta_T$$

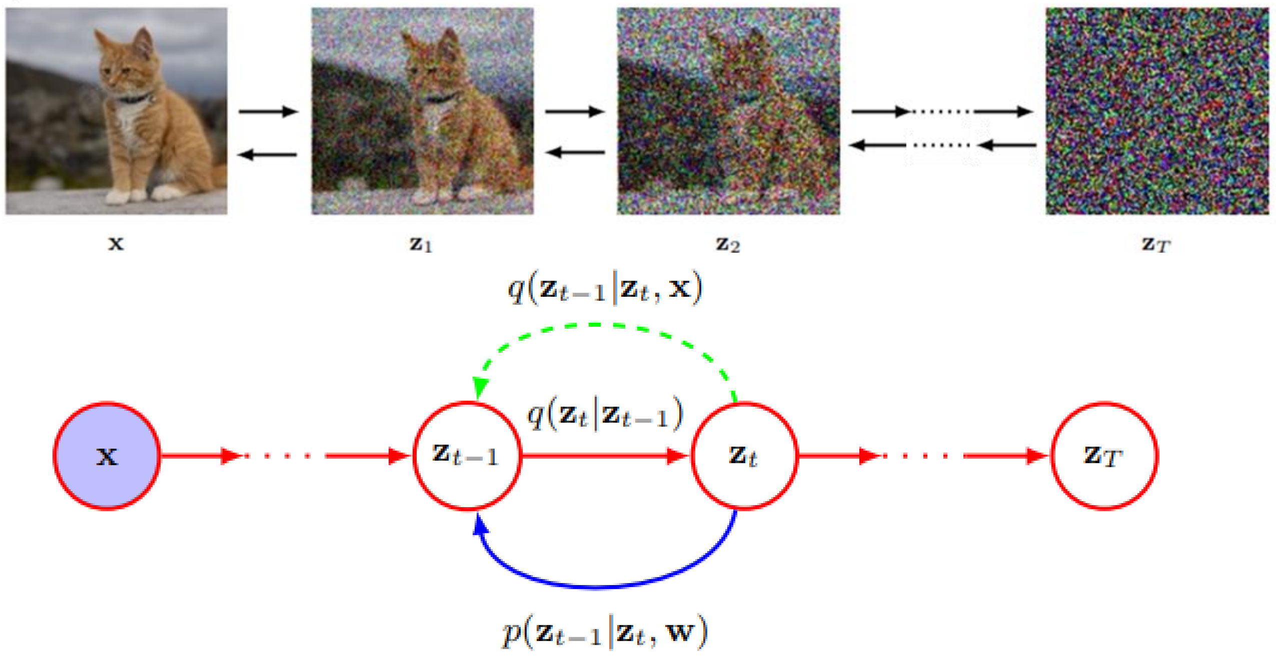
Increase each step

Both coefficients ensures that, for each step, mean and variance of z_t **gets closer 0 and to I**

For $T \to \infty$, $\beta_T \to 1$ and z_T would be a variable within the prior defined noise distribution $N(0, \mathbf{I})$

Diffusion-based Models

- Diffusion models [46] typically involve two stages:
  - **Forward encoder**
    - Gradually corrupts input data by adding noise to it
    - After several steps input data is transformed on a known noise-like distribution
  - **Reverse decoder**
    - Progressively denoise the data to reconstruct
    - p(x) is unknown!
    - Often parametrized by a Neural Network



Source: Bishop and Bishop [5]

$$\mathbf{z}_t = \sqrt{1-\beta_t}\mathbf{z}_{t-1} + \sqrt{\beta_t}\boldsymbol{\epsilon}_t$$

$$\beta_t \ll 1$$

- $\uparrow$ Inference time
- $\downarrow$ shift between $\hat{x}$ and x

$\uparrow T \rightarrow$

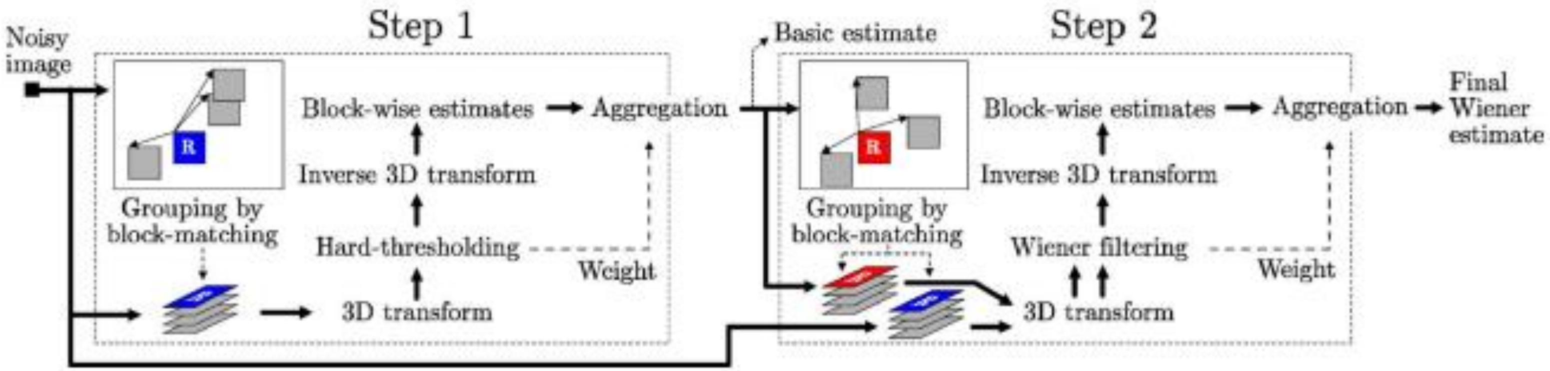
Reverting each step should be easier as the distance between $q(z_{t-1})$ and $q(z)$ distributions would be narrower

- $\downarrow$ Inference time
- $\uparrow$ shift between $\hat{x}$ and x

$\downarrow T \rightarrow$

- **Non-Learning Classical Methods**
  - Non-Local Means [8]
  - Block-Matching and 3D Filtering [13]

BM3D



Source: Dabov et al. [13]

Posters / **146**

Accelerating Neutron Tomography Ring Artifact Removal Using BM3DORNL

Author: Chen Zhang[1]

Co-authors: Dmitry Ganyushin [1]; Jose Borreguero-Calvo [1]; Pete Peterson [1]

[1] *Oak Ridge National Laboratory*

Neutron tomography is a crucial tool for material examination, but ring artifacts can significantly decrease data quality and complicate tasks like segmentation and morphological analysis. The Block-Matching and 3D filtering (BM3D) algorithm, known for mitigating vertical streaks in sinograms and addressing the root cause of ring artifacts, is unfortunately slow and CPU-intensive. We introduce a unique, open-source software solution that eliminates ring artifacts in neutron tomography using the BM3D algorithm. By leveraging both CPU acceleration through Numba and GPU acceleration through CuPy, our approach significantly improves computational efficiency while maintaining data integrity. This dual-acceleration framework drastically speeds up BM3D processing, allowing researchers to quickly obtain refined results and streamline segmentation and morphological analysis.

Abstract publication:

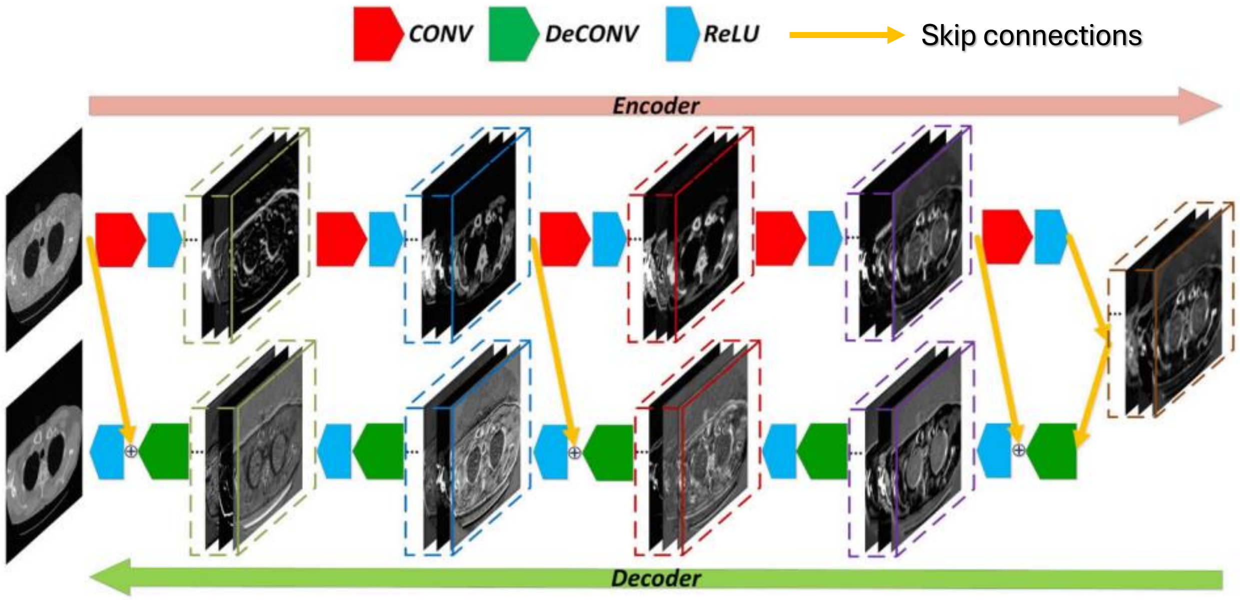
I agree that the abstract will be published on the web site

Source: NOBUGS 2024

- Non-Learning Classical Methods
  - Non-Local Means [8]
  - Block-Matching and 3D Filtering [13]

- **CNN-based Methods**
  - RED-CNN [12]

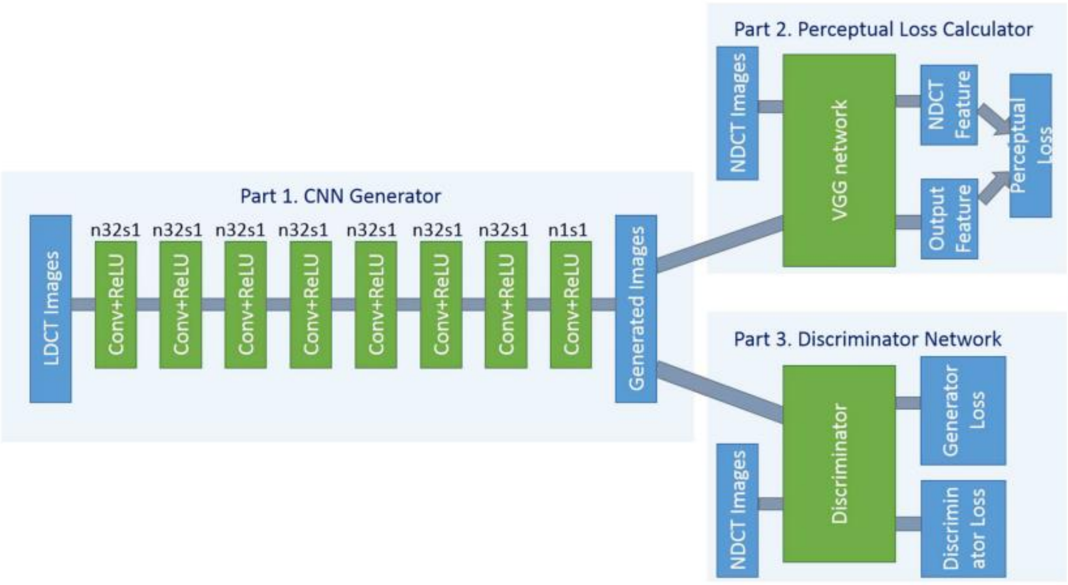
RED-CNN



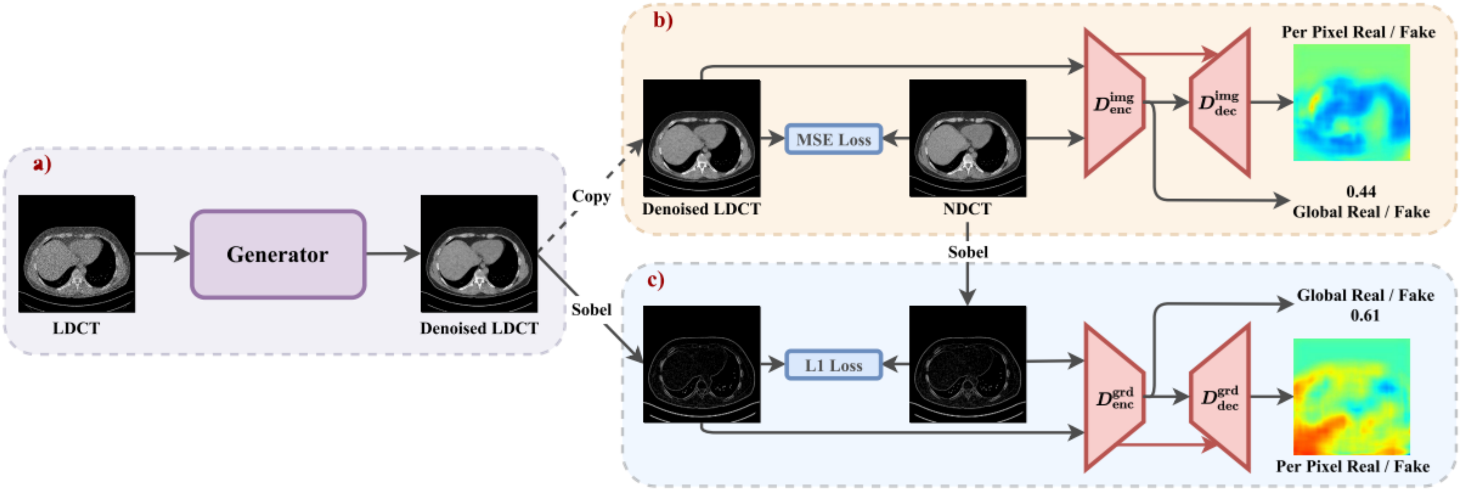
Source: Chen et al. [12]

- Non-Learning Classical Methods
  - Non-Local Means [8]
  - Block-Matching and 3D Filtering [13]
- CNN-based Methods
  - RED-CNN [12]
- **GAN-based Methods**
  - WGAN-VGG [57]
  - DU-GAN [27]

WGAN-VGG



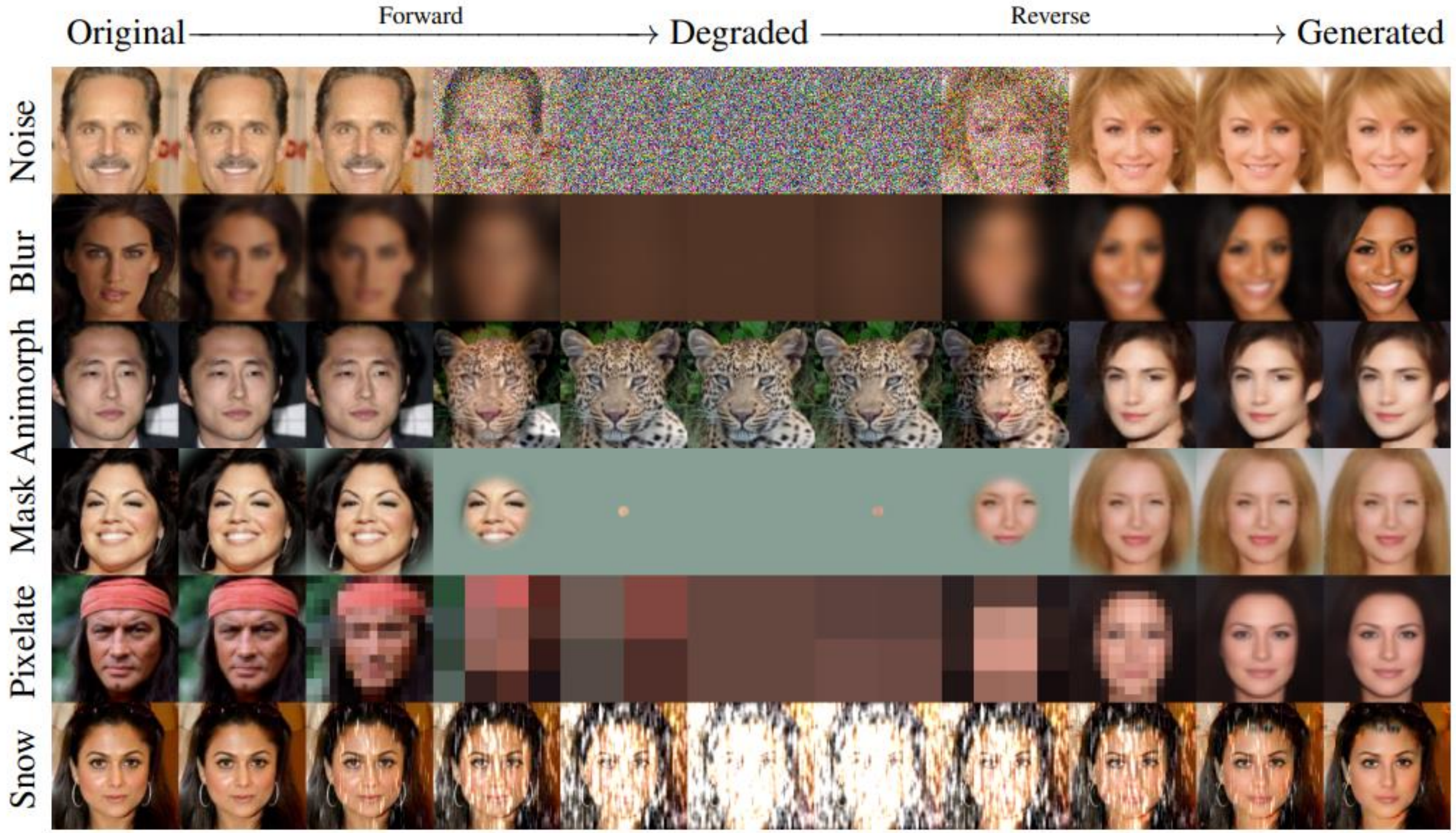
Source: Yang et al. [57]



Source: Huang et al. [27]

- Non-Learning Classical Methods
  - Non-Local Means [8]
  - Block-Matching and 3D Filtering [13]
- CNN-based Methods
  - RED-CNN [12]
- GAN-based Methods
  - WGAN-VGG [57]
  - DU-GAN [27]
- **Diffusion-based Methods**
  - Cold Diffusion [2]
  - DDPM [56]
  - Contextual Conditional Diffusion model (CoCoDiff) [18]
  - Denoising with Diffusion Prior (Dn-Dp) [35]
  - Contextual Error-modulated Generalized Diffusion Mode (CoreDiff) [17]

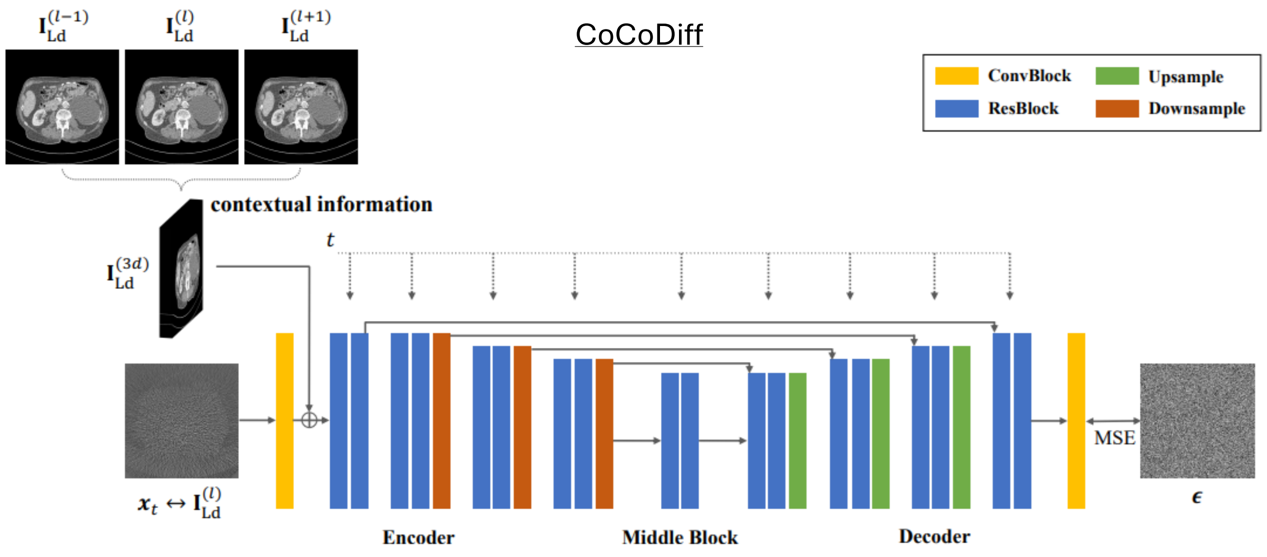
ColdDiffusion



Source: Bansal et al. [2]

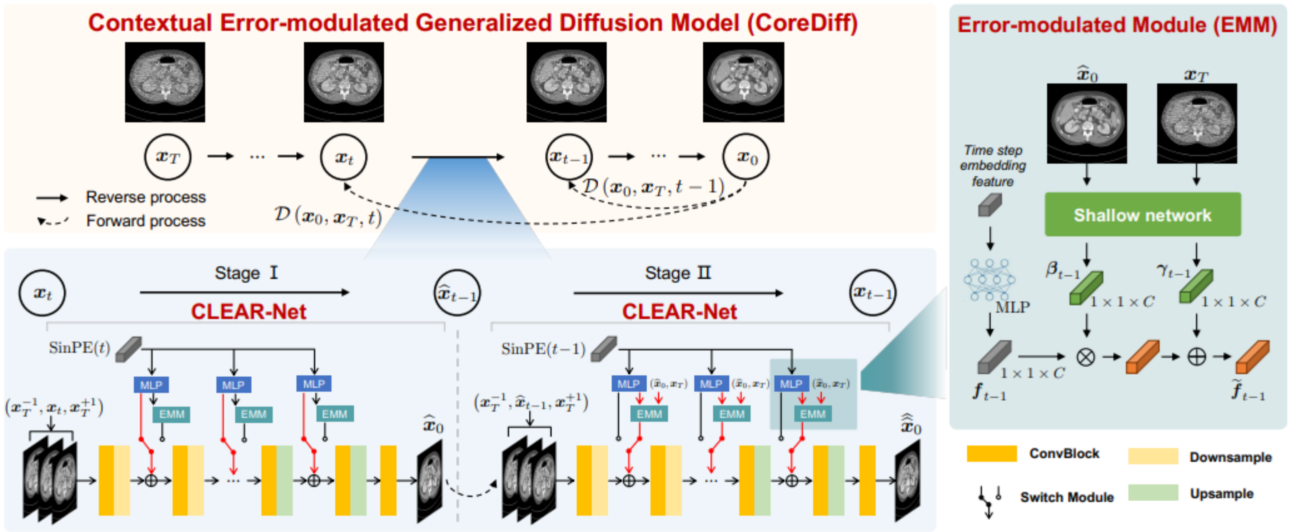
- Non-Learning Classical Methods
  - Non-Local Means [8]
  - Block-Matching and 3D Filtering [13]
- CNN-based Methods
  - RED-CNN [12]
- GAN-based Methods
  - WGAN-VGG [57]
  - DU-GAN [27]
- **Diffusion-based Methods**
  - Cold Diffusion [2]
  - DDPM [56]
  - Contextual Conditional Diffusion model (CoCoDiff) [18]
  - Denoising with Diffusion Prior (Dn-Dp) [35]
  - Contextual Error-modulated Generalized Diffusion Mode (CoreDiff) [17]

CoCoDiff

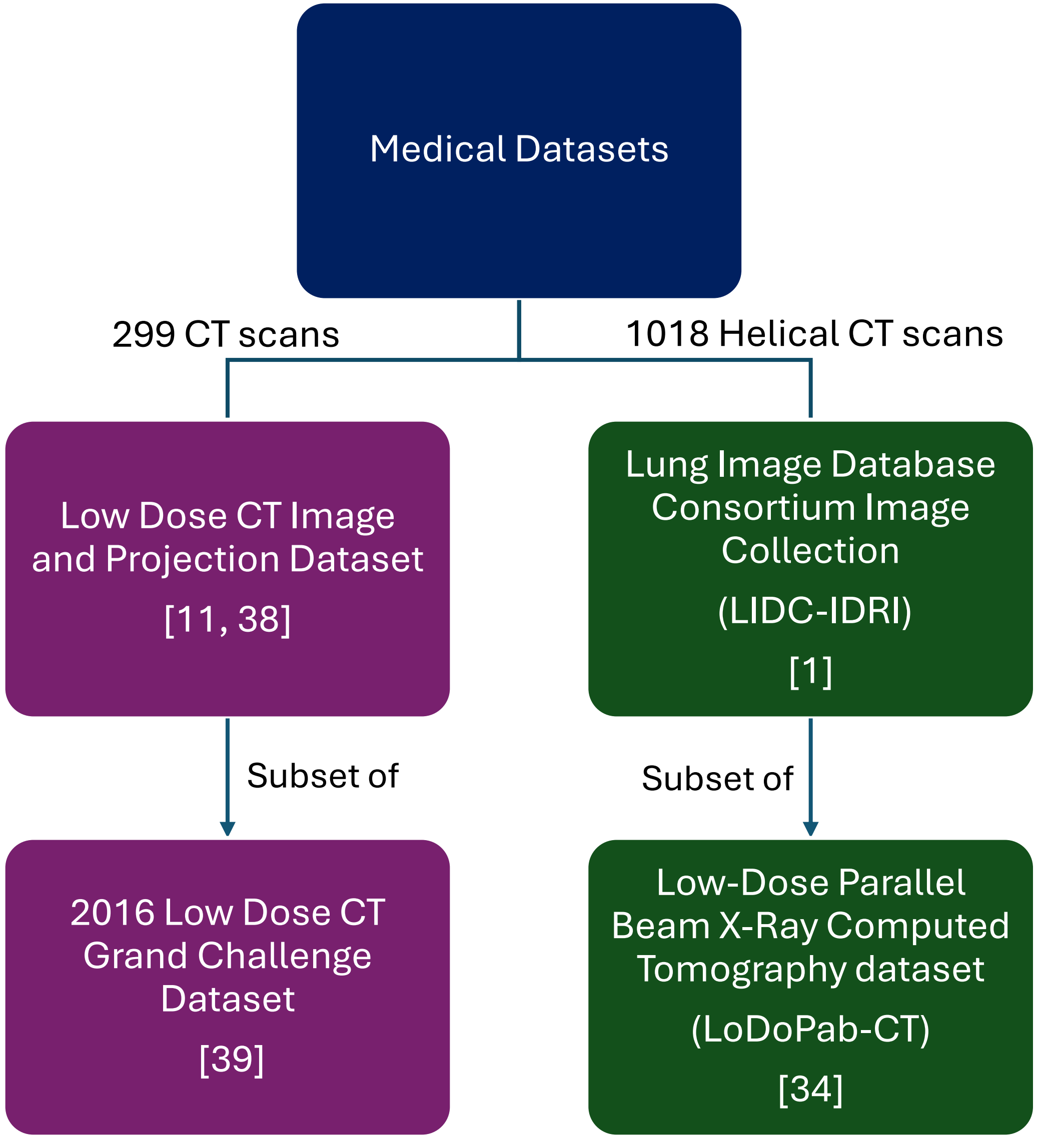


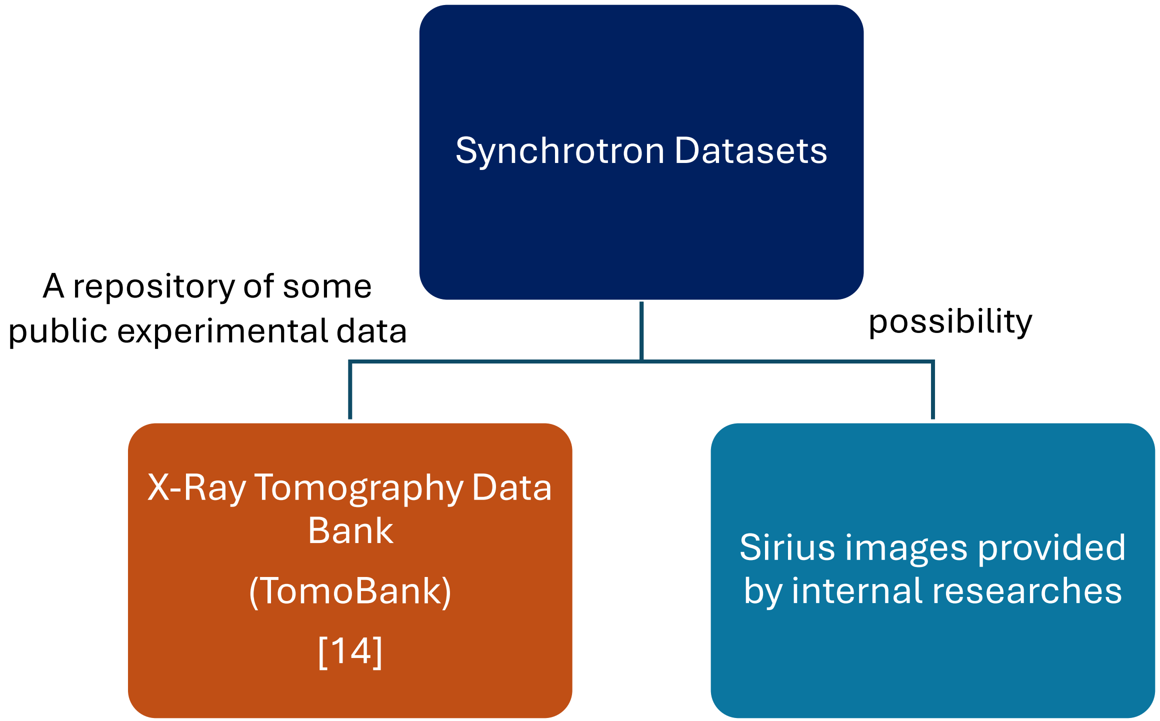
Source: Gao and Shan [18]

CoreDiff



Source: Gao et al. [17]





- **Number of Medical CT scans available is much higher than synchrotron CT scans!**
- **Low-dose data are simulated**

Root Mean Square Error [20]

$$\mathrm{MSE} = \frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \left(I_{ij} - \hat{I}_{ij} \right)^2$$

$$\mathrm{RMSE} = \sqrt{\mathrm{MSE}}$$

Peak Signal-to-Noise Ratio [25]

$$\mathrm{PSNR} = 10 \times \log_{10} \left(\frac{L_I^2}{\mathrm{MSE}} \right)$$

Structural Similarity Index Measure [52]

$$\mathrm{SSIM}(\mathbf{x}, \mathbf{y}) = [l(\mathbf{x}, \mathbf{y})]^{\alpha} \times [c(\mathbf{x}, \mathbf{y})]^{\beta} \times [s(\mathbf{x}, \mathbf{y})]^{\gamma}$$

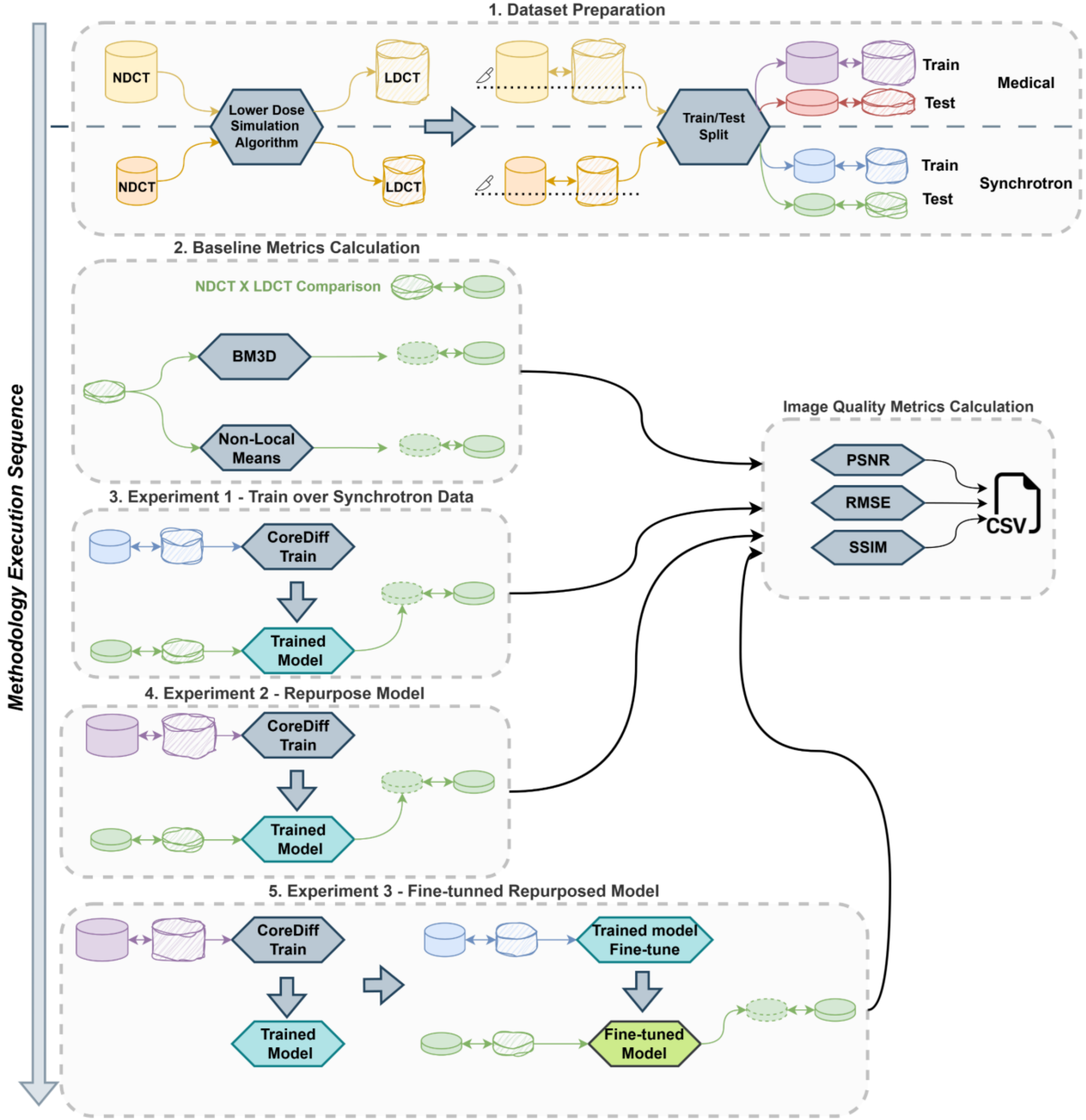
$$l(\mathbf{x}, \mathbf{y}) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \qquad C_1 = (k_1 L)^2$$

$$c(\mathbf{x}, \mathbf{y}) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \qquad C_2 = (k_2 L)^2$$

$$s(\mathbf{x}, \mathbf{y}) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \qquad C_3 = \frac{C_2}{2}$$

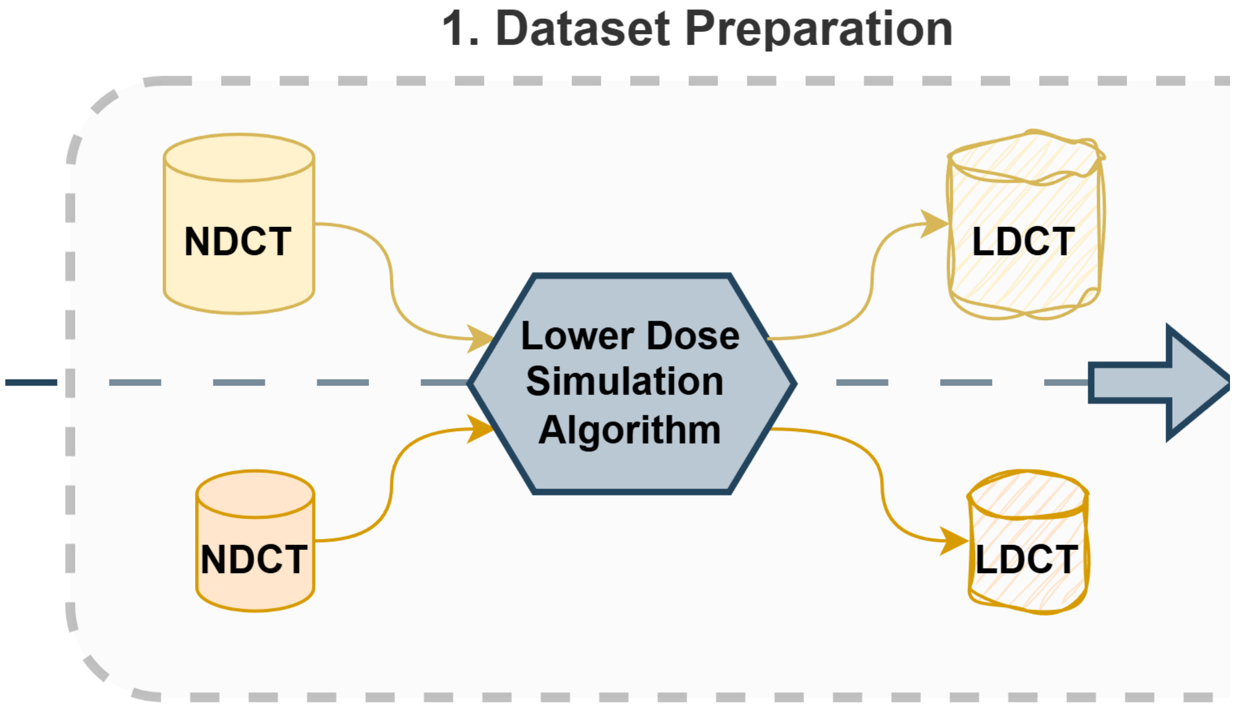
$$k_1 = 0.01 \quad k_2 = 0.03$$

- Methodology was divided into 5 stages:



- Methodology was divided into 5 stages:

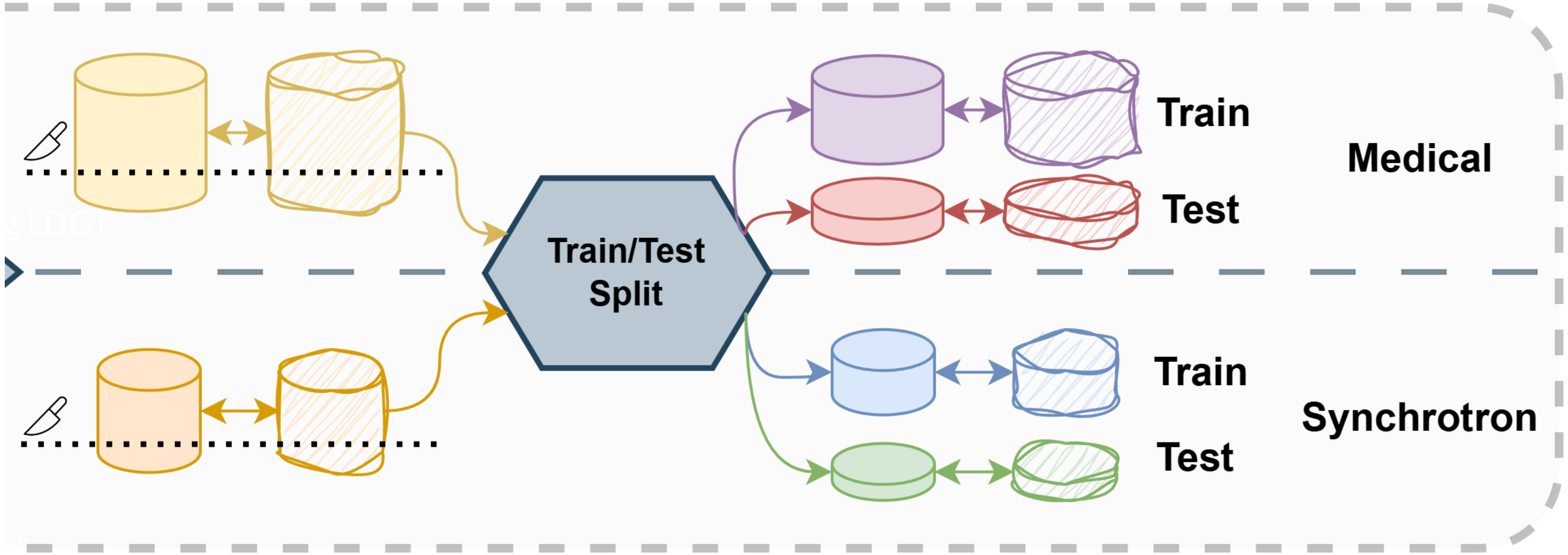
- **Stage 1: Dataset Preparation**
  - Simulate LDCT data for a range of different dose levels using the algorithm proposed by Yu et al. [58]



- Methodology was divided into 5 stages:

- **Stage 1: Dataset Preparation**
  - Simulate LDCT data for a range of different dose levels using the algorithm proposed by Yu et al. [58]
  - Train/Test set split

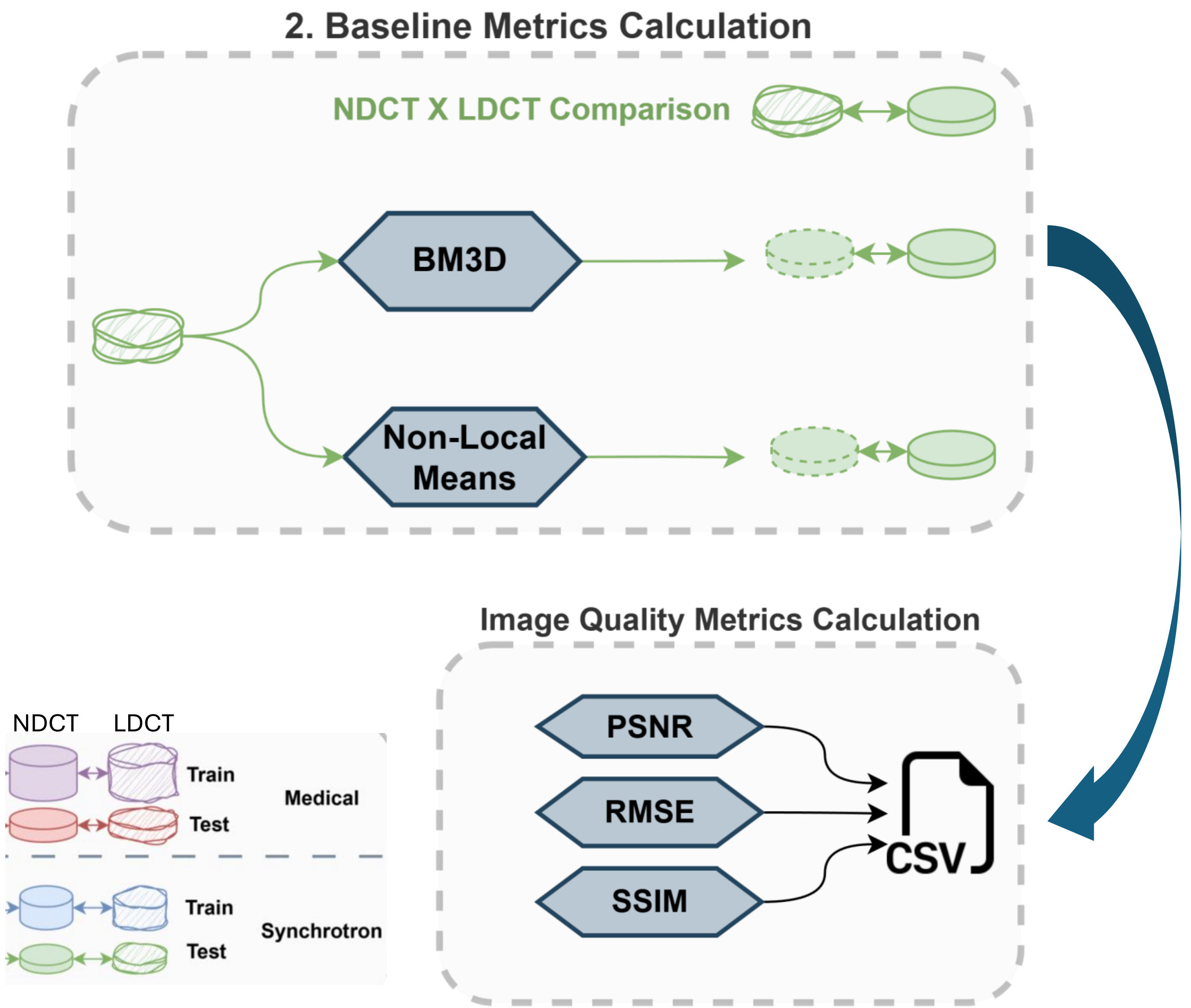
1. Dataset Preparation



- Methodology was divided into 5 stages:

- Stage 1: Dataset Preparation
  - Simulate LDCT data for a range of different dose levels using the algorithm proposed by Yu et al. [58]
  - Train/Test set split

- **Stage 2: Baseline Metrics Calculation**
  - Compare simulated LDCT and NDCT data for baseline metrics
  - Test BM3D and Non-Local Means for denoising LDCT synchrotron data

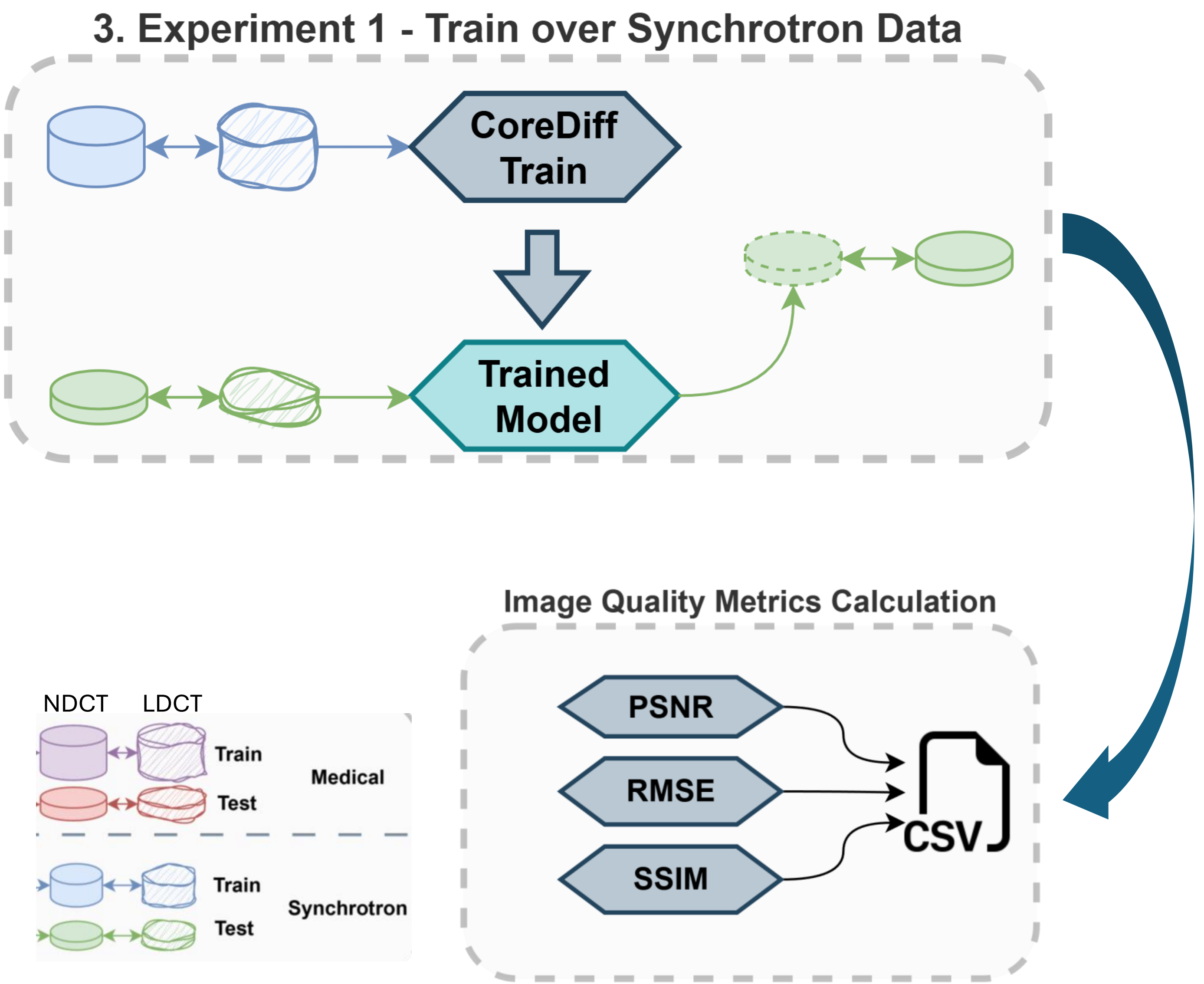


- Methodology was divided into 5 stages:

- Stage 1: Dataset Preparation
  - Simulate LDCT data for a range of different dose levels using the algorithm proposed by Yu et al. [58]
  - Train/Test set split

- Stage 2: Baseline Metrics Calculation
  - Compare simulated LDCT and NDCT data for baseline metrics
  - Test BM3D and Non-Local Means for denoising LDCT synchrotron data

- **Stage 3: Experiment 1**
  - Train and Test CoreDiff over synchrotron data



- Methodology was divided into 5 stages:

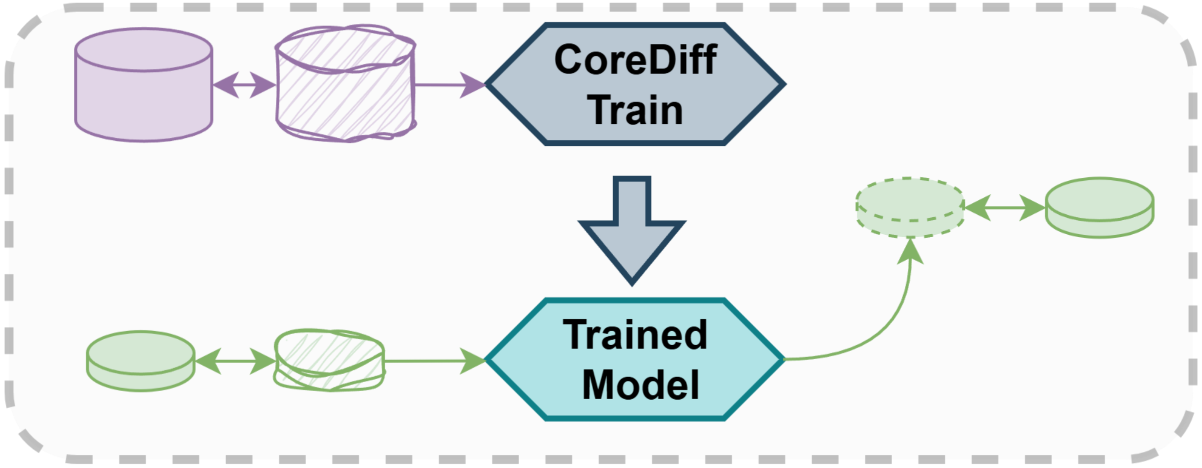
- Stage 1: Dataset Preparation
  - Simulate LDCT data for a range of different dose levels using the algorithm proposed by Yu et al. [58]
  - Train/Test set split

- Stage 2: Baseline Metrics Calculation
  - Compare simulated LDCT and NDCT data for baseline metrics
  - Test BM3D and Non-Local Means for denoising LDCT synchrotron data

- Stage 3: Experiment 1
  - Train and Test CoreDiff over synchrotron data

- **Stage 4: Experiment 2**
  - Train CoreDiff over Medical data and test it over synchrotron data

4. Experiment 2 - Repurpose Model



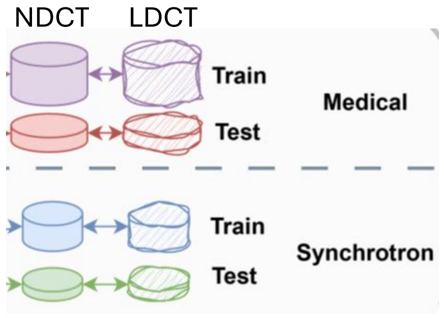
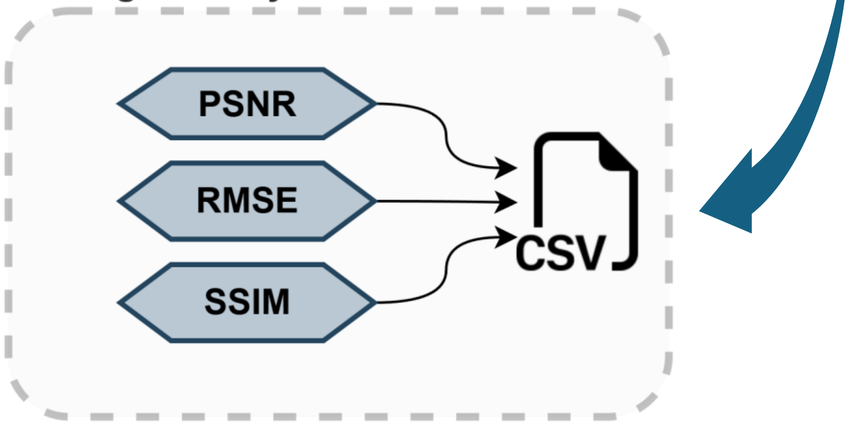


Image Quality Metrics Calculation



- Methodology was divided into 5 stages:

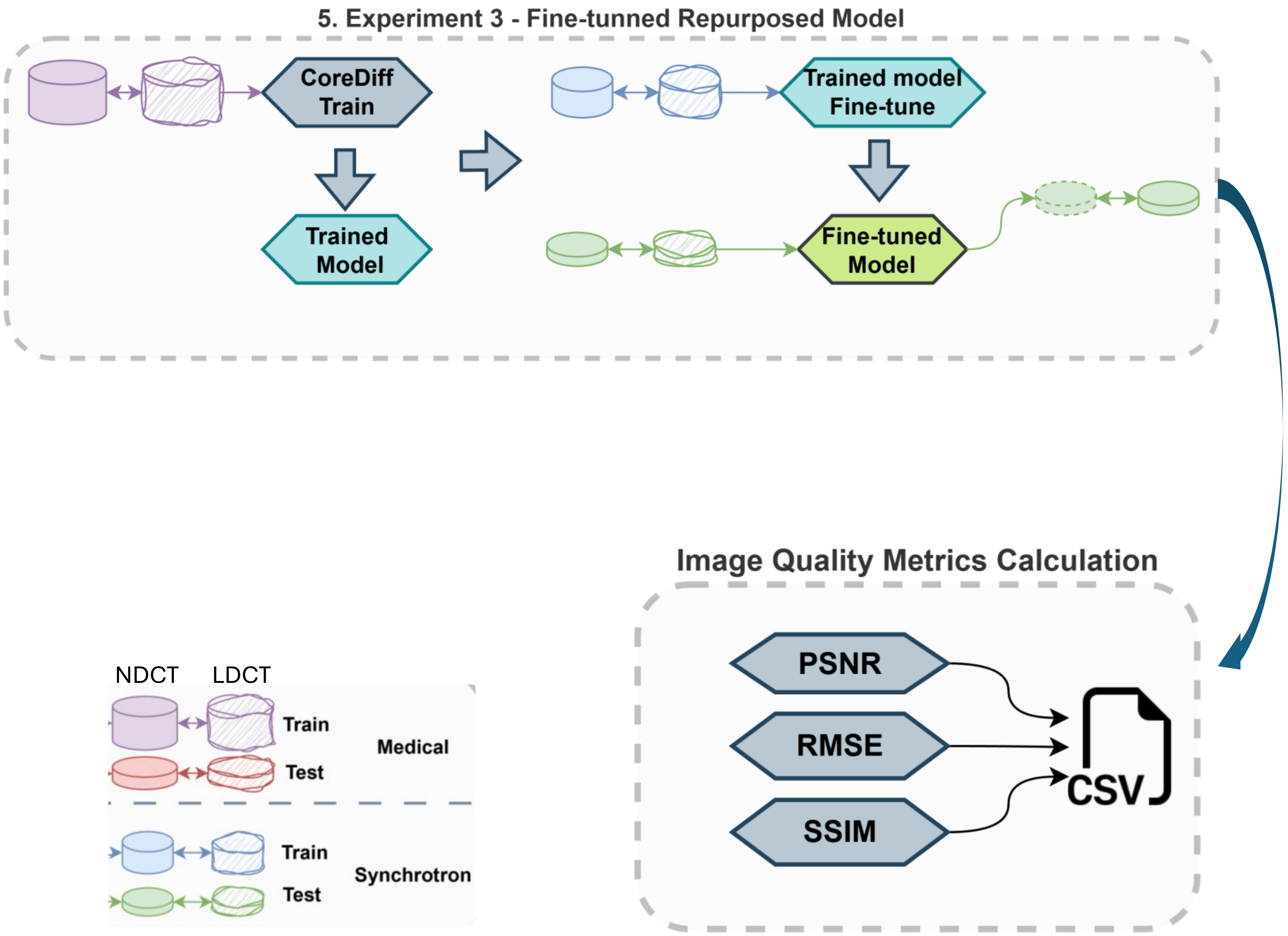
- Stage 1: Dataset Preparation
  - Simulate LDCT data for a range of different dose levels using the algorithm proposed by Yu et al. [58]
  - Train/Test set split

- Stage 2: Baseline Metrics Calculation
  - Compare simulated LDCT and NDCT data for baseline metrics
  - Test BM3D and Non-Local Means for denoising LDCT synchrotron data

- Stage 3: Experiment 1
  - Train and Test CoreDiff over synchrotron data

- Stage 4: Experiment 2
  - Train CoreDiff over Medical data and test it over synchrotron data

- **Stage 5: Experiment 3**
  - Train CoreDiff over Medical data, finetune it using synchrotron data and test it over synchrotron



A. Literature review

B. Proposal writing and realization of the qualification exam

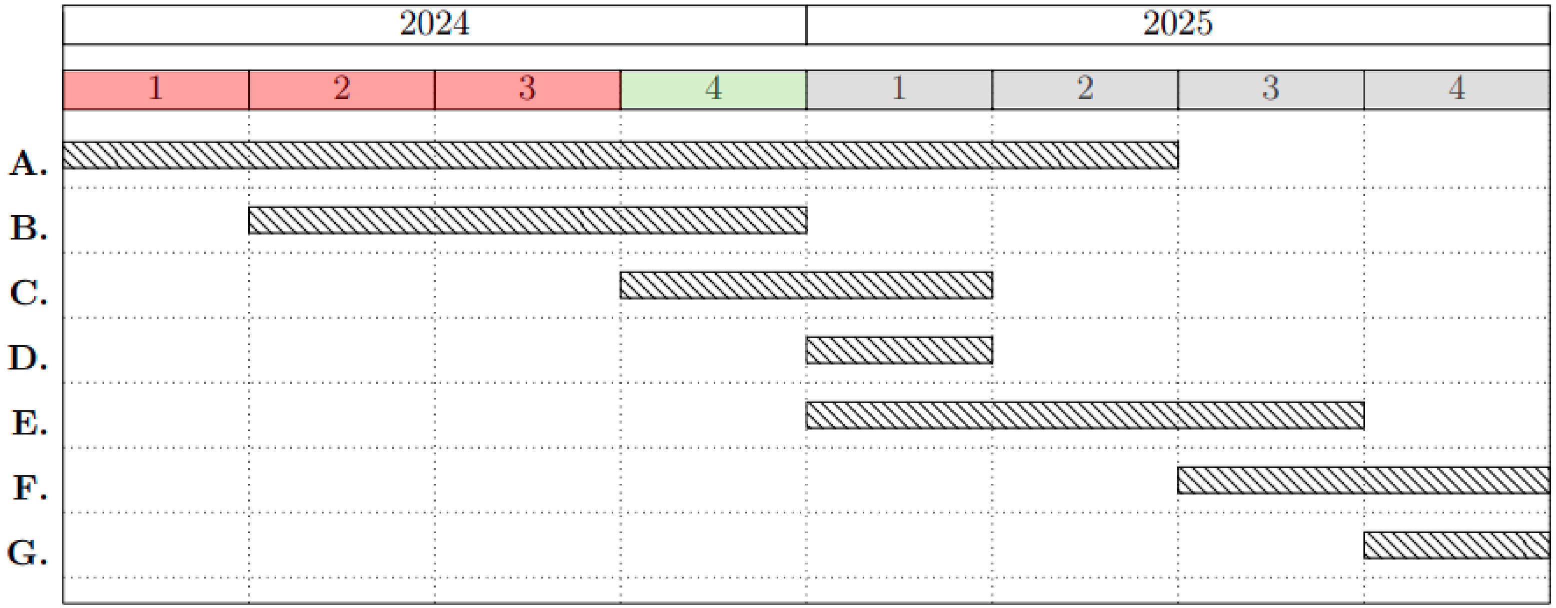
C. Dataset preparation

D. Baseline metric calculation

E. Experiments execution and results comparison

F. Results documentation and publishing

G. Dissertation writing and defense



References

- * POLO, Carla C. et al. Correlations between lignin content and structural robustness in plants revealed by X-ray ptychography. **Scientific reports**, v. 10, n. 1, p. 6023, 2020.
- ** BUSHONG, Stewart C. Radiologic science for technologists. 1988.
- *** NAKAZATO, Muka; ITO, Sosuke. Geometrical aspects of entropy production in stochastic thermodynamics based on Wasserstein distance. **Physical Review Research**, v. 3, n. 4, p. 043093, 2021.

- [1] S. G. Armato III, G. McLennan, L. Bidaut, M. F. McNitt-Gray, C. R. Meyer, A. P. Reeves, B. Zhao, D. R. Aberle, C. I. Henschke, E. A. Hoffman, E. A. Kazerooni, H. MacMahon, E. J. R. Van Beek, D. Yankelevitz, A. M. Biancardi, P. H. Bland, M. S. Brown, R. M. Engelmann, G. E. Laderach, D. Max, R. C. Pais, D. P. Y. Qing, R. Y. Roberts, A. R. Smith, A. Starkey, P. Batra, P. Caligiuri, A. Farooqi, G. W. Gladish, C. M. Jude, R. F. Munden, I. Petkovska, L. E. Quint, L. H. Schwartz, B. Sundaram, L. E. Dodd, C. Fenimore, D. Gur, N. Petrick, J. Freymann, J. Kirby, B. Hughes, A. V. Casteele, S. Gupte, M. Sallam, M. D. Heath, M. H. Kuhn, E. Dharaiya, R. Burns, D. S. Fryd, M. Salganicoff, V. Anand, U. Shreter, S. Vastagh, B. Y. Croft, and L. P. Clarke. Data from The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A completed reference database of lung nodules on CT scans, 2015. The Cancer Imaging Archive. https://doi.org/10.7937/K9/TCIA.2015.LO9QL9SX.
- [2] A. Bansal, E. Borgnia, H.-M. Chu, J. Li, H. Kazemi, F. Huang, M. Goldblum, J. Geiping, and T. Goldstein. Cold Diffusion: Inverting Arbitrary Image Transforms Without Noise. In Advances in Neural Information Processing Systems 36 (NeurIPS), volume 36, pages 41259–41282, 2023.
- [3] D. Bhatt, C. Patel, H. Talsania, J. Patel, R. Vaghela, S. Pandya, K. Modi, and H. Ghayvat. CNN Variants for Computer Vision: History, Architecture, Application, Challenges and Future Scope. Electronics, 10(20):2470, 2021.
- [4] C. M. Bishop. Latent Variable Models. In Learning in Graphical Models, pages 371– 403. Springer, 1998.
- [5] C. M. Bishop and H. Bishop. Deep Learning: Foundations and Concepts. Springer Nature, 2023.
- [6] F. E. Boas and D. Fleischmann. CT artifacts: causes and reduction techniques. Imaging in Medicine, 4(2):229–240, 2012.
- [7] U. Bonse and F. Busch. X-Ray computed microtomography (μCT) using synchrotron radiation (SR). Progress in Biophysics and Molecular Biology, 65(1):133–169, 1996.
- [8] A. Buades, B. Coll, and J.-M. Morel. A non-local algorithm for image denoising. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), volume 2, pages 60–65 vol. 2, 2005.
- [9] J. T. Bushberg, J. A. Seibert, J. Leidholdt, Edwin M, and J. M. Boone. The essential physics of medical imaging. Lippincott Williams & Wilkins, 2011.

References

- [10] B. Chen, X. Duan, Z. Yu, S. Leng, L. Yu, and C. McCollough. Development and Validation of an Open Data Format for CT Projection Data. Medical Physics, 42(12):6964– 6972, 2015.
- [11] B. Chen, S. Leng, L. Yu, D. H. III, J. Fletcher, and C. McCollough. An open library of CT patient projection data. In Medical Imaging 2016: Physics of Medical Imaging, volume 9783, page 97831B. International Society for Optics and Photonics, SPIE, 2016.
- [12] H. Chen, Y. Zhang, M. K. Kalra, F. Lin, Y. Chen, P. Liao, J. Zhou, and G. Wang. Low-Dose CT With a Residual Encoder-Decoder Convolutional Neural Network. IEEE Transactions on Medical Imaging, 36(12):2524–2535, 2017.
- [13] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering. IEEE Transactions on Image Processing, 16(8):2080–2095, 2007.
- [14] F. De Carlo, D. Gürsoy, D. J. Ching, K. J. Batenburg, W. Ludwig, L. Mancini, F. Marone, R. Mokso, D. M. Pelt, J. Sijbers, and M. Rivers. TomoBank: a tomographic data repository for computational X-Ray science. Measurement Science and Technology, 29(3):034004, 2018.
- [15] X. Duan, X. F. Ding, N. Li, F.-X. Wu, X. Chen, and N. Zhu. Sparse2Noise: Low-dose synchrotron X-Ray tomography without high-quality reference data. Computers in Biology and Medicine, 165:107473, 2023.
- [16] X. Duan, N. Li, X. Chen, and N. Zhu. Characterization of Tissue Scaffolds Using Synchrotron Radiation Microcomputed Tomography Imaging. Tissue Engineering Part C: Methods, 27(11):573–588, 2021.
- [17] Q. Gao, Z. Li, J. Zhang, Y. Zhang, and H. Shan. CoreDiff: Contextual ErrorModulated Generalized Diffusion Model for Low-Dose CT Denoising and Generalization. IEEE Transactions on Medical Imaging, 43(2):745–759, 2024.
- [18] Q. Gao and H. Shan. CoCoDiff: a contextual conditional diffusion model for low-dose CT image denoising. In Developments in X-Ray Tomography XIV, volume 12242, page 122420I. International Society for Optics and Photonics, SPIE, 2022.
- [19] K. Gong, K. Johnson, G. El Fakhri, Q. Li, and T. Pan. PET image denoising based on denoising diffusion probabilistic model. European Journal of Nuclear Medicine and Molecular Imaging, 51(2):358–368, 2024.
- [20] R. Gonzalez and R. Woods. Digital Image Processing. Prentice Hall, 2008.
- [21] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative Adversarial Nets. Advances in Neural Information Processing Systems 27 (NeurIPS), 27, 2014.
- [22] Google Colab. Goggle Colaboratory. https: // colab. google/ , 2024. Online; accessed on August 6, 2024.
- [23] C. A. Hamm, H. Mallison, O. Hampe, D. Schwarz, J. Mews, J. Blobel, A. S. Issever, and P. Asbach. Efficiency, workflow and image quality of clinical computed tomography scanning compared to photogrammetry on the example of a Tyrannosaurus rex skull from the Maastrichtian of Montana, USA. Journal of Paleontological Techniques, 21:1– 13, 2018.

References

- [24] J. Ho, A. Jain, and P. Abbeel. Denoising diffusion probabilistic models. Advances in Neural Information Processing Systems 33 (NeurIPS), 33:6840–6851, 2020.
- [25] A. Hore and D. Ziou. Image Quality Metrics: PSNR vs. SSIM. In 20th International Conference on Pattern Recognition, pages 2366–2369. IEEE, 2010.
- [26] D. Hu, Y. K. Tao, and I. Oguz. Unsupervised denoising of retinal OCT with diffusion probabilistic model. In Medical Imaging 2022: Image Processing, volume 12032, pages 25–34. SPIE, 2022.
- [27] Z. Huang, J. Zhang, Y. Zhang, and H. Shan. DU-GAN: Generative Adversarial Networks With Dual-Domain U-Net-Based Discriminators for Low-Dose CT Denoising. IEEE Transactions on Instrumentation and Measurement, 71:1–12, 2022.
- [28] V.-P. Karjalainen, M. A. Finnilä, P. L. Salmon, and S. Lipkin. Micro-computed tomography imaging and segmentation of the archaeological textiles from Valmarinniemi. Journal of Archaeological Science, 160:105871, 2023.
- [29] A. Kazerouni, E. K. Aghdam, M. Heidari, R. Azad, M. Fayyaz, I. Hacihaliloglu, and D. Merhof. Diffusion models in medical imaging: A comprehensive survey. Medical Image Analysis, 88:102846, 2023.
- [30] D. Kingma, T. Salimans, B. Poole, and J. Ho. Variational Diffusion Models. Advances in Neural Information Processing Systems 34 (NeurIPS), 34:21696–21707, 2021.
- [31] K. S. H. Kulathilake, N. A. Abdullah, A. Q. M. Sabri, and K. W. Lai. A review on deep learning approaches for low-dose computed tomography restoration. Complex & Intelligent Systems, 9(3):2713–2745, 2023.
- [32] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. Nature, 521(7553):436–444, 2015.
- [33] LeCun, Yann and Kavukcuoglu, Koray and Farabet, Clement. Convolutional Networks and Applications in Vision. In International Symposium on Circuits and Systems, pages 253–256. IEEE, 2010.
- [34] J. Leuschner, M. Schmidt, D. O. Baguer, and P. Maass. LoDoPaB-CT, a benchmark dataset for low-dose computed tomography reconstruction. Scientific Data, 8(1):109, 2021.
- [35] X. Liu, Y. Xie, S. Diao, S. Tan, and X. Liang. A diffusion probabilistic prior for zero-shot low-dose CT image denoising. arXiv preprint arXiv:2305.15887, 2023.
- [36] Z. Liu, T. Bicer, R. Kettimuthu, D. Gursoy, F. D. Carlo, and I. Foster. TomoGAN: lowdose synchrotron X-Ray tomography with generative adversarial networks: discussion. Journal of the Optical Society of America A, 37(3):422–434, 2020.
- [37] Martin Arjovsky and Soumith Chintala and Léon Bottou. Wasserstein Generative Adversarial Networks. In International Conference on Machine Learning (ICML), volume 70 of Proceedings of Machine Learning Research, pages 214–223, 2017.
- [38] C. McCollough, B. Chen, D. R. Holmes III, X. Duan, Z. Yu, L. Yu, S. Leng, and J. Fletcher. Low Dose CT Image and Projection Data, 2020. The Cancer Imaging Archive. https://doi.org/10.7937/9NPB-2637.

References

- [39] C. H. McCollough, A. C. Bartley, R. E. Carter, B. Chen, T. A. Drees, P. Edwards, D. R. Holmes III, A. E. Huang, F. Khan, S. Leng, K. L. McMillan, G. J. Michalak, K. M. Nunez, L. Yu, and J. G. Fletcher. Low-dose CT for the Detection and Classification of Metastatic Liver Lesions: Results of the 2016 Low Dose CT Grand Challenge. Medical Physics, 44(10):e339–e352, 2017.
- [40] E. Okuno and E. M. Yoshimura. Física das Radiações. Oficina de Textos, 2016.
- [41] F. Rosenblatt. Principles of neurodynamics: Perceptrons and the theory of brain mechanisms, volume 55. Spartan books Washington, D.C., 1962.
- [42] S. V. M. Sagheer and S. N. George. A review on medical image denoising algorithms. Biomedical Signal Processing and Control, 61:102036, 2020.
- [43] P. Savoia, S. K. Jayanthi, and M. C. Chammas. Focused Assessment with Sonography for Trauma (FAST). Journal of Medical Ultrasound, 31(2):101–106, 2023
- [44] R. Schilling, B. Jastram, O. Wings, D. Schwarz-Wings, and A. S. Issever. Reviving the dinosaur: virtual reconstruction and three-dimensional printing of a dinosaur vertebra. Radiology, 270(3):864–871, 2014.
- [45] K. Simonyan and A. Zisserman. Very Deep Convolutional Networks for Large-scale Image Recognition. arXiv preprint arXiv:1409.1556, 2014.
- [46] J. Sohl-Dickstein, E. Weiss, N. Maheswaranathan, and S. Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. In International Conference on Machine Learning (ICML), volume 37 of Proceedings of Machine Learning Research, pages 2256–2265. PMLR, 2015.
- [47] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole. Scorebased generative modeling through stochastic differential equations. arXiv preprint arXiv:2011.13456, 2020.
- [48] Y. R. Tonin. Coherent X-Ray Diffraction Imaging: Image reconstruction via a matrix model of the inhomogenous Helmholtz equation. Master's thesis, Unicamp, 2022.
- [49] N. T. Trung, D.-H. Trinh, N. L. Trung, and M. Luong. Low-dose CT image denoising using deep convolutional neural networks with extended receptive fields. Signal, Image and Video Processing, 16(7):1963–1971, 2022.
- [50] L. Vásárhelyi, Z. Kónya, A. Kukovecz, and R. Vajtai. Microcomputed tomography–based characterization of advanced materials: a review. Materials Today Advances, 8:100084, 2020.
- [51] T. Wang, Y. Lei, Z. Tian, X. Dong, Y. Liu, X. Jiang, W. J. Curran, T. Liu, H.- K. Shu, and X. Yang. Deep learning-based image quality improvement for low-dose computed tomography simulation in radiation therapy. Journal of Medical Imaging, 6(4):043504–043504, 2019.

References

- [52] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image Quality Assessment: From Error Visibility to Structural Similarity. IEEE Transactions on Image Processing, 13(4):600–612, 2004.
- [53] Z. Wang, E. P. Simoncelli, and A. C. Bovik. Multiscale Structural Similarity for Image Quality Assessment. In Asilomar Conference on Signals, Systems & Computers (ACSSC), volume 2, pages 1398–1402. IEEE, 2003.
- [54] L. Weng. What are diffusion models? https: // lilianweng. github. io/ posts/ 2021-07-11-diffusion-models , 2021. Online; accessed on July 8, 2024.
- [55] P. Willmott. An introduction to synchrotron radiation: techniques and applications. John Wiley & Sons, 2019.
- [56] W. Xia, Q. Lyu, and G. Wang. Low-Dose CT Using Denoising Diffusion Probabilistic Model for 20× Speedup. arXiv preprint arXiv:2209.15136, 2022.
- [57] Q. Yang, P. Yan, Y. Zhang, H. Yu, Y. Shi, X. Mou, M. K. Kalra, Y. Zhang, L. Sun, and G. Wang. Low-Dose CT Image Denoising Using a Generative Adversarial Network With Wasserstein Distance and Perceptual Loss. IEEE Transactions on Medical Imaging, 37(6):1348–1357, 2018.
- [58] L. Yu, M. Shiung, D. Jondal, and C. H. McCollough. Development and Validation of a Practical Lower-Dose-Simulation Tool for Optimizing Computed Tomography Scan Protocols. Journal of Computer Assisted Tomography, 36(4):477–487, 2012.

Enhancing Synchrotron Low-Dose Computed Tomography Image Quality Using Diffusion-Based Generative Models

Candidate: Paulo Baraldi Mausbach

Supervisor: Prof. Dr. Zanoni Dias

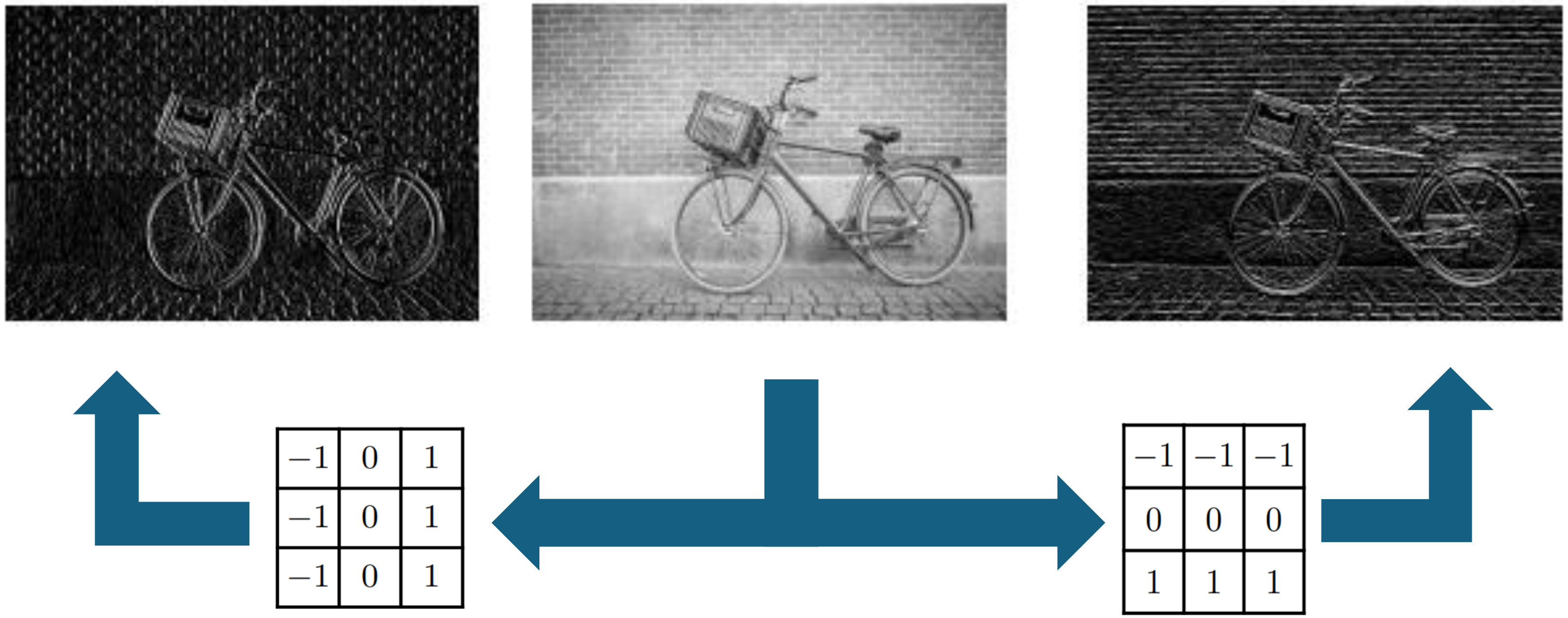
Co-supervisor: Prof. Dr. Hélio Pedrini

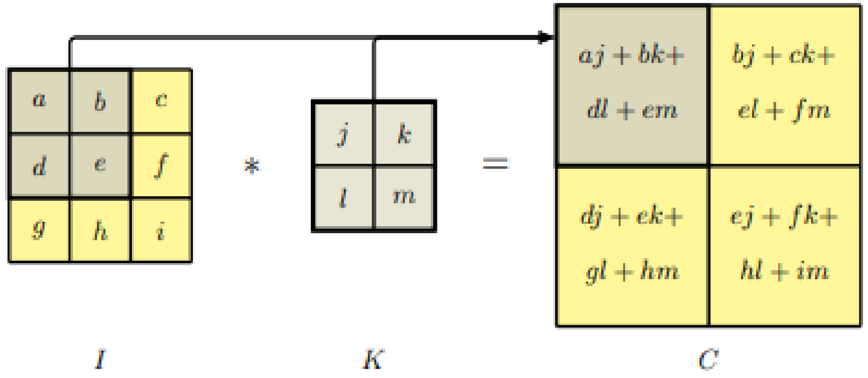
Master's Degree Proposal

Institute of Computing - University of Campinas

October 17th of 2024

- Classical approaches:
  - Extensive feature extraction work
  - Handcrafted filters for edge, texture, shape
  - Some approaches required pre-processing data to work
- Convolutional Neural Networks (CNNs):
  - Processing done directly on raw data
  - Learn filters based on data
  - Learn simple and complex filters that optimize target task
- CNN revolutionized Computer Vision (CV) area [5]
- It is also extensively used for other data dimension applications such as time series, 3D images and videos [5]





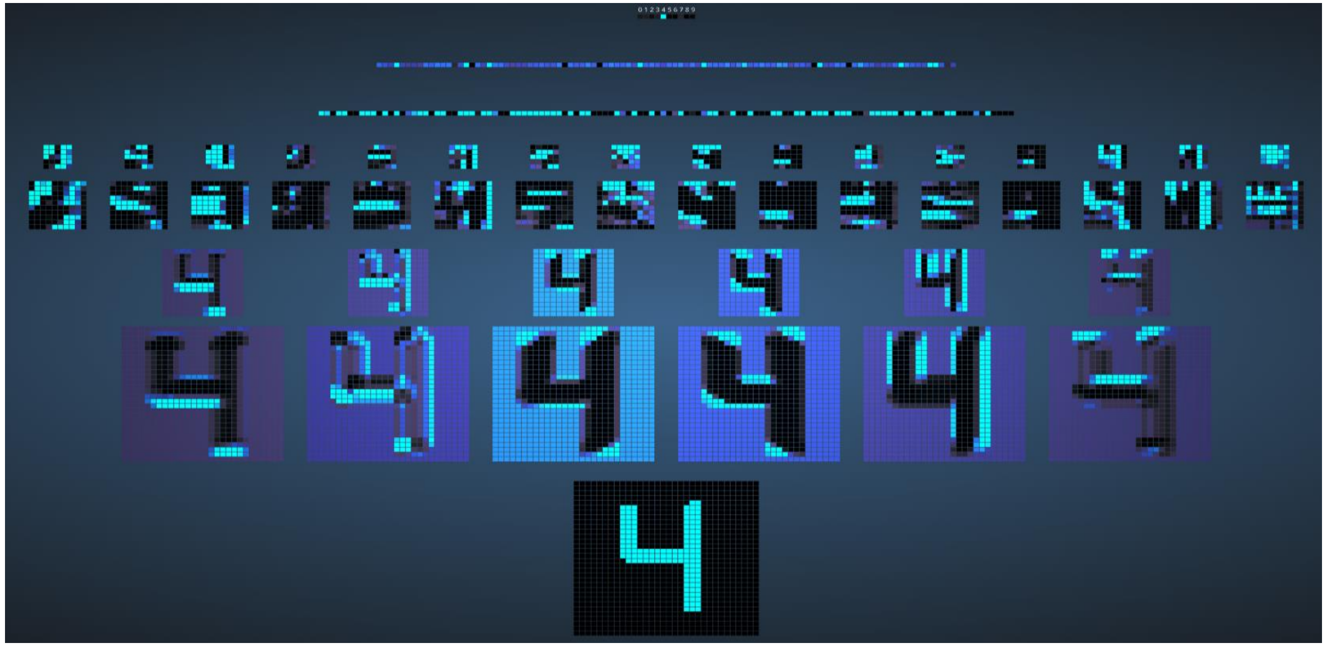
*Images Source: Bishop and Bishop [5]

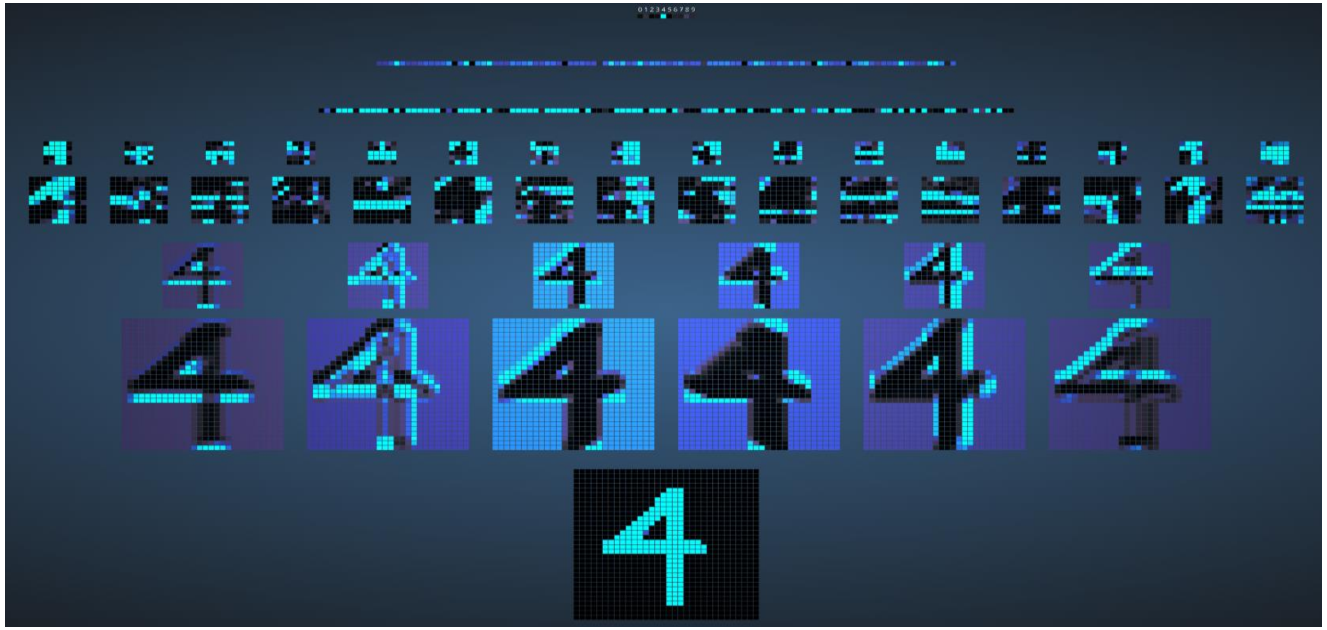
- Classical approaches:
  - Extensive feature extraction work
  - Handcrafted filters for edge, texture, shape
  - Some approaches required pre-processing data to work

- Convolutional Neural Networks (CNNs):
  - Processing done directly on raw data
  - Learn filters based on data
  - Learn simple and complex filters that optimize target task

- CNN revolutionized Computer Vision (CV) area [5]

- It is also extensively used for other data dimension applications such as time series, 3D images and vídeos [5]





Source: https://adamharley.com/nn_vis/cnn/2d.html

Root Mean Square Error (RMSE)

- ↓RMSE (Root Mean Square Error) [20]
- Difference between all pixel values from the "noise-free" image and the denoised estimation
- Low RMSE value indicates high similarity between "noise-free" and denoised estimation image

Image height Image width

$$\mathrm{MSE} = \frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \left(I_{ij} - \hat{I}_{ij} \right)^2$$

Pixel value from position (i,j) in the noise-free image

Pixel value from position (i,j) in the denoised estimation

$$\mathrm{RMSE} = \sqrt{\mathrm{MSE}}$$

- ↑Peak Signal-to-Noise Ratio [25]

- Commonly used for assessing quality of image reconstruction, compression and denoising algorithms

- Ratio between the maximum possible signal value ($L_{\hat{I}}$) and the noise corrupting it (MSE between I and $\hat{I}$)

- $MSE \rightarrow 0$, $PSNR \rightarrow \infty$

- High PSNR indicates low degradation of the signal by the existing noise

Maximum pixel value from noise-free image

$$\mathrm{PSNR} = 10 \times \log_{10} \left(\frac{L_I^2}{\mathrm{MSE}} \right)$$

$$\mathrm{MSE} = \frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \left(I_{ij} - \hat{I}_{ij} \right)^2$$

Image height

Image width

Pixel value from position (i,j) in the noise-free image

Pixel value from position (i,j) in the denoised estimation

Structural Similarity Index Measure (SSIM)

- SSIM (Structural Similarity Index Measure) [52]
- Perception-based metric to measure similarity between two images
- Calculated by comparing the degradation between two same size windows taken from the same position in the "noise-free" and the denoised estimation image.
- Evaluates three image aspects:
  - $l(\boldsymbol{x}, \boldsymbol{y}) \rightarrow$ Luminance
  - $c(\boldsymbol{x}, \boldsymbol{y}) \rightarrow$ Contrast
  - $s(\boldsymbol{x}, \boldsymbol{y}) \rightarrow$ Structure
- Calculated as a weighted combination of all three aspects

$$\mathrm{SSIM}(\mathbf{x}, \mathbf{y}) = [l(\mathbf{x}, \mathbf{y})]^{\alpha} \times [c(\mathbf{x}, \mathbf{y})]^{\beta} \times [s(\mathbf{x}, \mathbf{y})]^{\gamma}$$

Weight factors

Mean pixel intensity

$$l(\mathbf{x}, \mathbf{y}) = \frac{2\mu_x \mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}$$

Standard deviation

$$c(\mathbf{x}, \mathbf{y}) = \frac{2\sigma_x \sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}$$

Variance

Covariance

$$s(\mathbf{x}, \mathbf{y}) = \frac{\sigma_{xy} + C_3}{\sigma_x \sigma_y + C_3}$$

- SSIM (Structural Similarity Index Measure) [52]
- Perception-based metric to measure similarity between two images
- Calculated by comparing the degradation between two same size windows taken from the same position in the "noise-free" and the denoised estimation image.
- Evaluates three image aspects:
    - $l(\boldsymbol{x},\boldsymbol{y})$ → Luminance
    - $c(\boldsymbol{x},\boldsymbol{y})$ → Contrast
    - $s(\boldsymbol{x},\boldsymbol{y})$ → Structure
- Calculated as a weighted combination of all three aspects

$$\mathrm{SSIM}(\mathbf{x},\mathbf{y}) = [l(\mathbf{x},\mathbf{y})]^{\alpha} \times [c(\mathbf{x},\mathbf{y})]^{\beta} \times [s(\mathbf{x},\mathbf{y})]^{\gamma}$$

Weight factors

Mean pixel intensity

$$l(\mathbf{x},\mathbf{y}) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}$$

Standard deviation

$$c(\mathbf{x},\mathbf{y}) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}$$

Variance

Covariance

$$s(\mathbf{x},\mathbf{y}) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}$$

Stabilization Variables

$$C_1 = (k_1L)^2$$

$$C_2 = (k_2L)^2$$

$$C_3 = \frac{C_2}{2}$$

- SSIM (Structural Similarity Index Measure) [52]
- Perception-based metric to measure similarity between two images
- Calculated by comparing the degradation between two same size windows taken from the same position in the "noise-free" and the denoised estimation image.
- Evaluates three image aspects:
  - $l(\boldsymbol{x}, \boldsymbol{y})$ → Luminance
  - $c(\boldsymbol{x}, \boldsymbol{y})$ → Contrast
  - $s(\boldsymbol{x}, \boldsymbol{y})$ → Structure
- Calculated as a weighted combination of all three aspects

$$\mathrm{SSIM}(\mathbf{x},\mathbf{y}) = [l(\mathbf{x},\mathbf{y})]^{\alpha} \times [c(\mathbf{x},\mathbf{y})]^{\beta} \times [s(\mathbf{x},\mathbf{y})]^{\gamma}$$

Weight factors

Mean pixel intensity

$$l(\mathbf{x},\mathbf{y}) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}$$

Standard deviation

$$c(\mathbf{x},\mathbf{y}) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}$$

Variance

Covariance

$$s(\mathbf{x},\mathbf{y}) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}$$

Stabilization Variables

$$C_1 = (k_1 L)^2$$

$$C_2 = (k_2 L)^2$$

$$C_3 = \frac{C_2}{2}$$

Maximum pixel value

$k_1 = 0.01 \quad k_2 = 0.03$

Small constants ($\ll 1$) with arbitrary value

SSIM authors reported "fairly insensitive" variations of SSIM caused by variations on values from k_1 and k_2 [52]

Structural Similarity Index Measure (SSIM)

- SSIM (Structural Similarity Index Measure) [52]
- Perception-based metric to measure similarity between two images
- Calculated by comparing the degradation between two same size windows taken from the same position in the “noise-free” and the denoised estimation image.
- Evaluates three image aspects:
  - $l(\boldsymbol{x},\boldsymbol{y}) \rightarrow$ Luminance
  - $c(\boldsymbol{x},\boldsymbol{y}) \rightarrow$ Contrast
  - $s(\boldsymbol{x},\boldsymbol{y}) \rightarrow$ Structure
- Calculated as a weighted combination of all three aspects

$$\mathrm{SSIM}(\mathbf{x},\mathbf{y}) = [l(\mathbf{x},\mathbf{y})]^{\alpha} \times [c(\mathbf{x},\mathbf{y})]^{\beta} \times [s(\mathbf{x},\mathbf{y})]^{\gamma}$$

SSIM can be simplified as $\xrightarrow{\text{if}}$ $\alpha = \beta = \gamma = 1$

$$\mathrm{SSIM}(\mathbf{x},\mathbf{y}) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

$$C_1 = (k_1 L)^2 \quad C_2 = (k_2 L)^2 \quad C_3 = \frac{C_2}{2}$$

$k_1 = 0.01 \quad k_2 = 0.03$