Models Genesis: Generic Autodidactic Models for 3D Medical Image Analysis

Zongwei Zhou[1], Vatsal Sodha[1], Md Mahfuzur Rahman Siddiquee[1], Ruibin Feng[1], Nima Tajbakhsh[1], Michael B. Gotway[2], and Jianming Liang[1]

[1] Arizona State University, Scottsdale, AZ 85259 USA
{zongweiz,vasodha,mrahmans,rfeng12,ntajbakh,jianming.liang}@asu.edu
[2] Mayo Clinic, Scottsdale, AZ 85259 USA
Gotway.Michael@mayo.edu

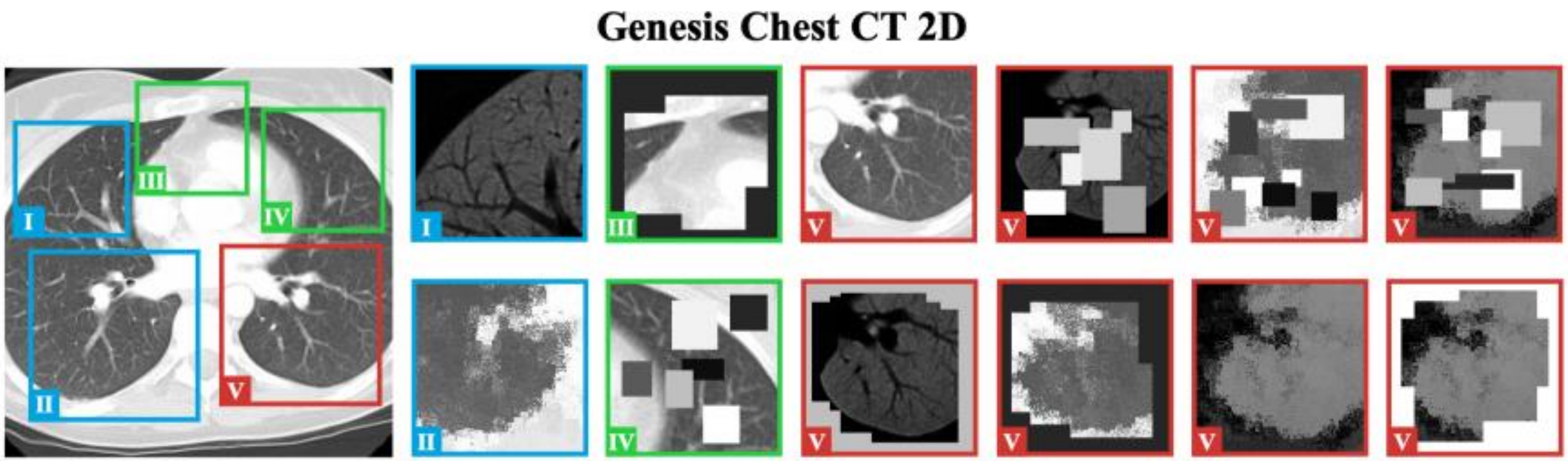
Presented by: Hengtao Guo
04/08/2020

Motivations

1. Transfer learning from natural image to medical image has been a practical paradigm.
2. However, 3D imaging tasks (like CT and MRI) have to sacrifice rich 3D context information and reformulated in 2D shape.
3. Learning a model from scratch simply in 3D may not necessarily yield performance better than transfer learning from ImageNet in 2D

Model Genesis

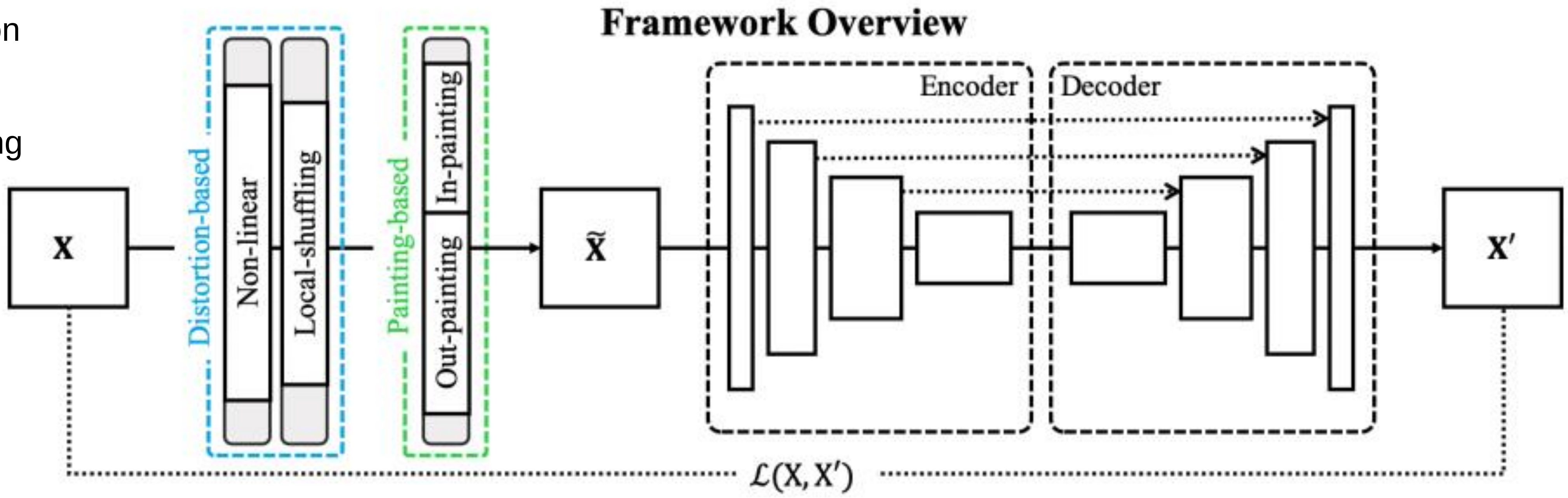
1. Hypothesis: Transfer learning can yield more powerful (application-specific) target models if the source models are built directly from medical images
2. Develops a framework that trains generic source models for 3D imaging (adaptive to different modalities and downstream tasks)
3. Model Genesis: models trained under this framework
   a) Without manual labeling
   b) Learned by self-supervision
   c) Served as source models for generating application-specific target models
4. Model Genesis can yield a common visual representation that is generalizable and transferable across diseases, organs and modalities.



Non-linear transformation

Local pixel shuffling

Out-painting & In-painting

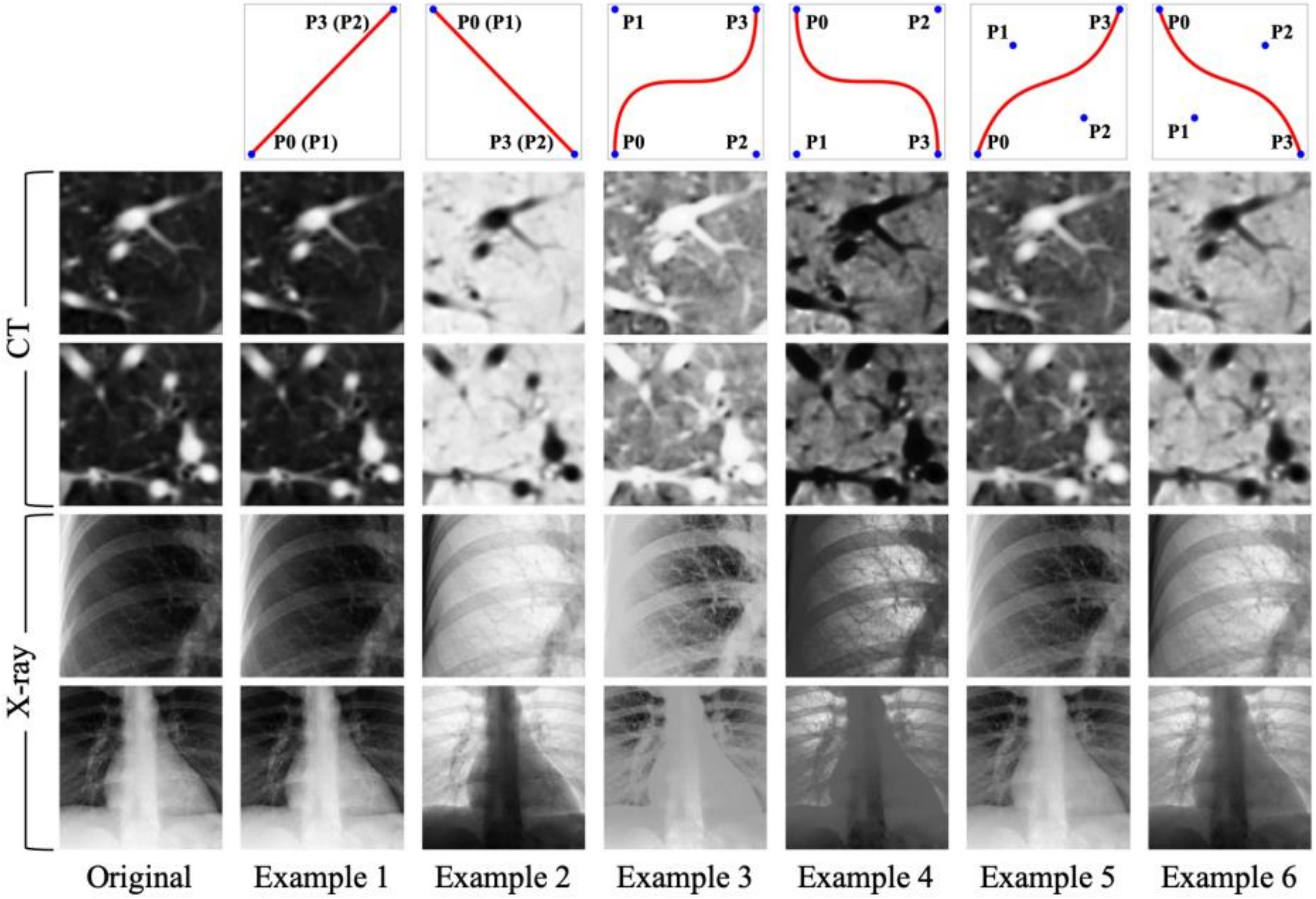


Transformations

1. Non-linear transformation
2. Local pixel shuffling
3. Out-painting & In-painting

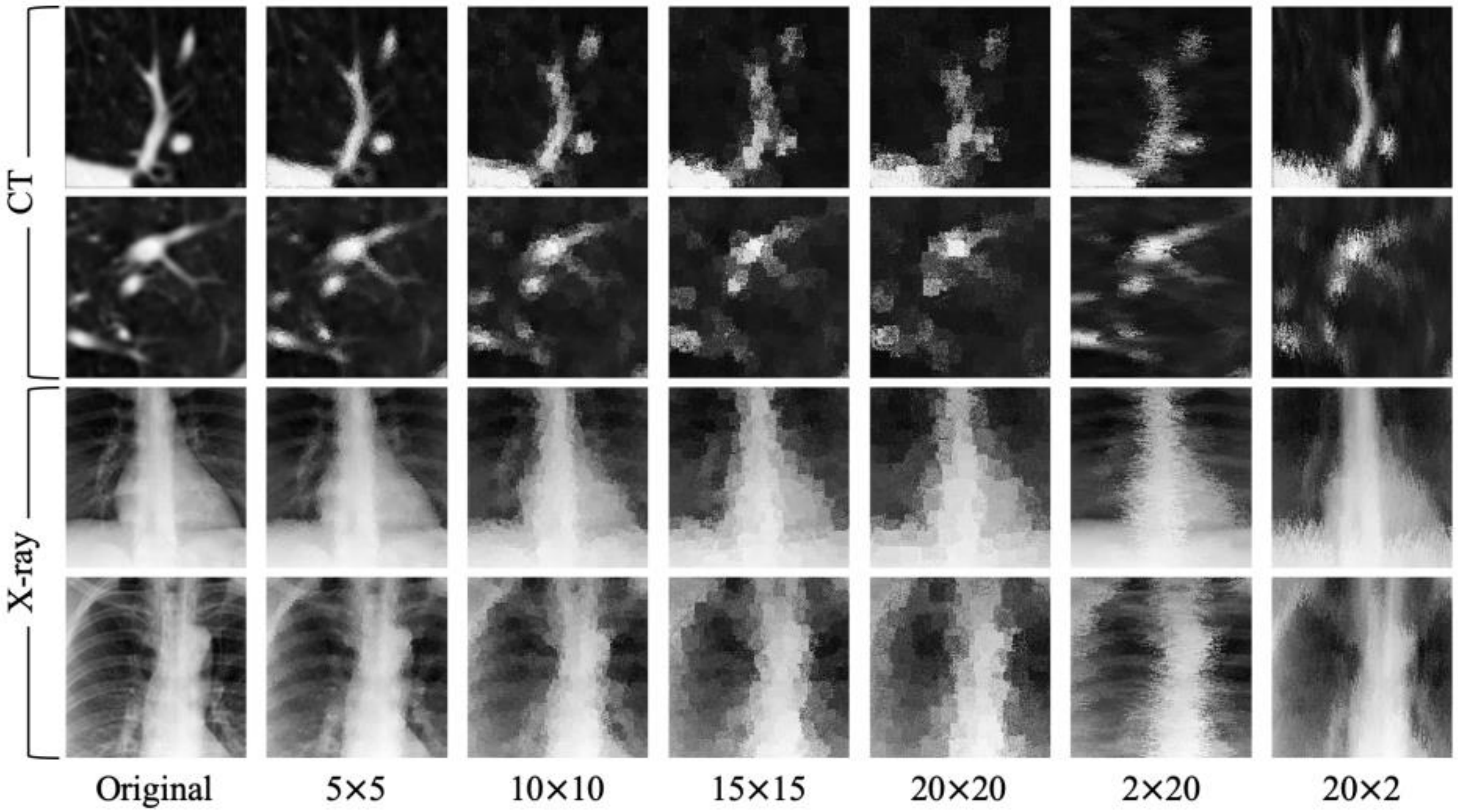
Non-linear transformation

1. Absolute or relative intensity values in medical images convey important information about the imaged structures and organs.
2. B´ezier Curve, a smooth and monotonous transformation function, which assigns every pixel a unique value.
3. Focuses Models Genesis on learning organ appearance (shape and intensity distribution)


P3 (P2)
P0 (P1)
P0 (P1)
P3 (P2)
P1
P3
P0
P2
P0
P2
P1
P3
P1
P3
P0
P2
P0
P2
P1
P3
CT
X-ray
Original
Example 1
Example 2
Example 3
Example 4
Example 5
Example 6

Local pixel shuffling

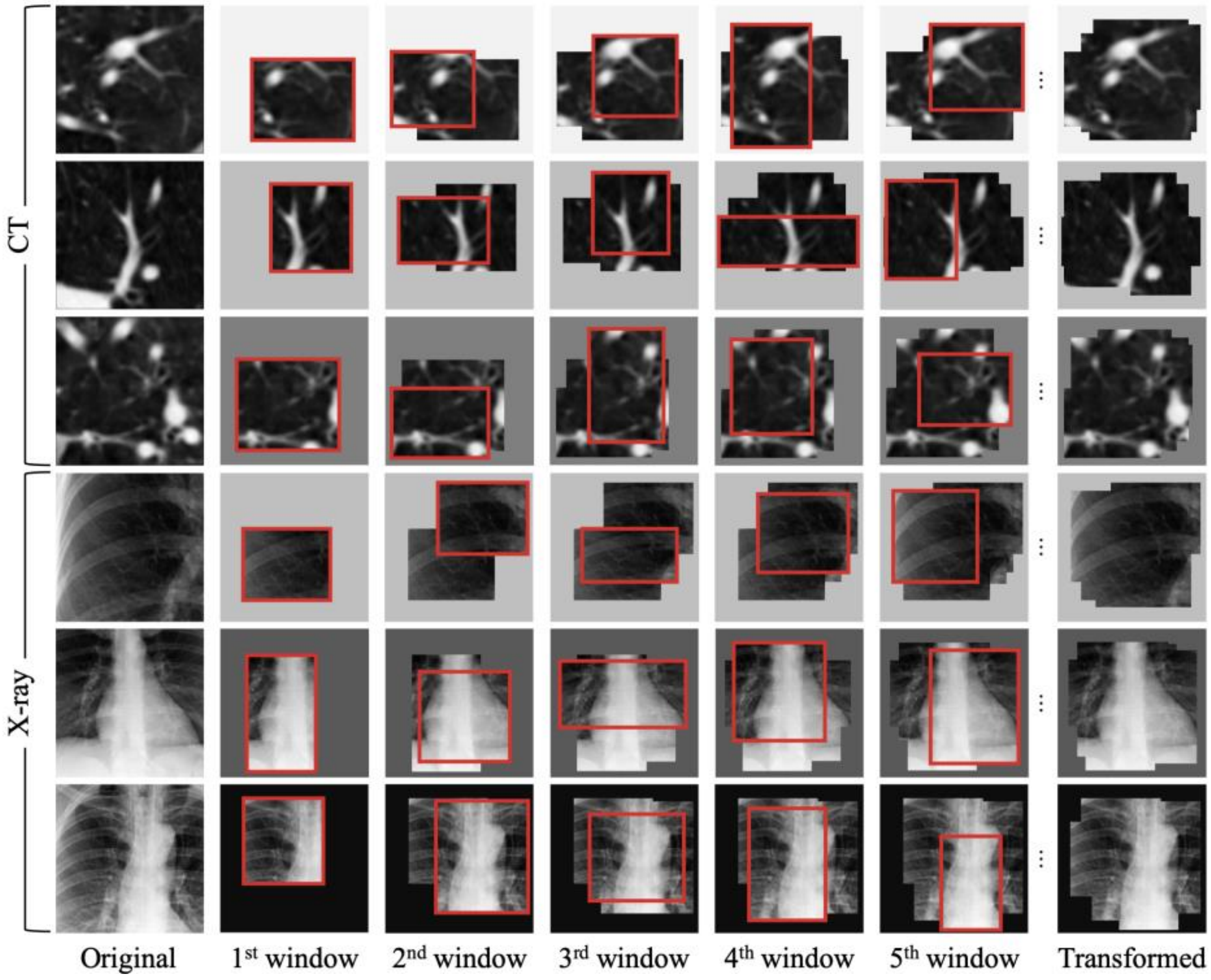
1. Randomly shuffle the pixels inside a local window, which is no bigger than the model's receptive field.



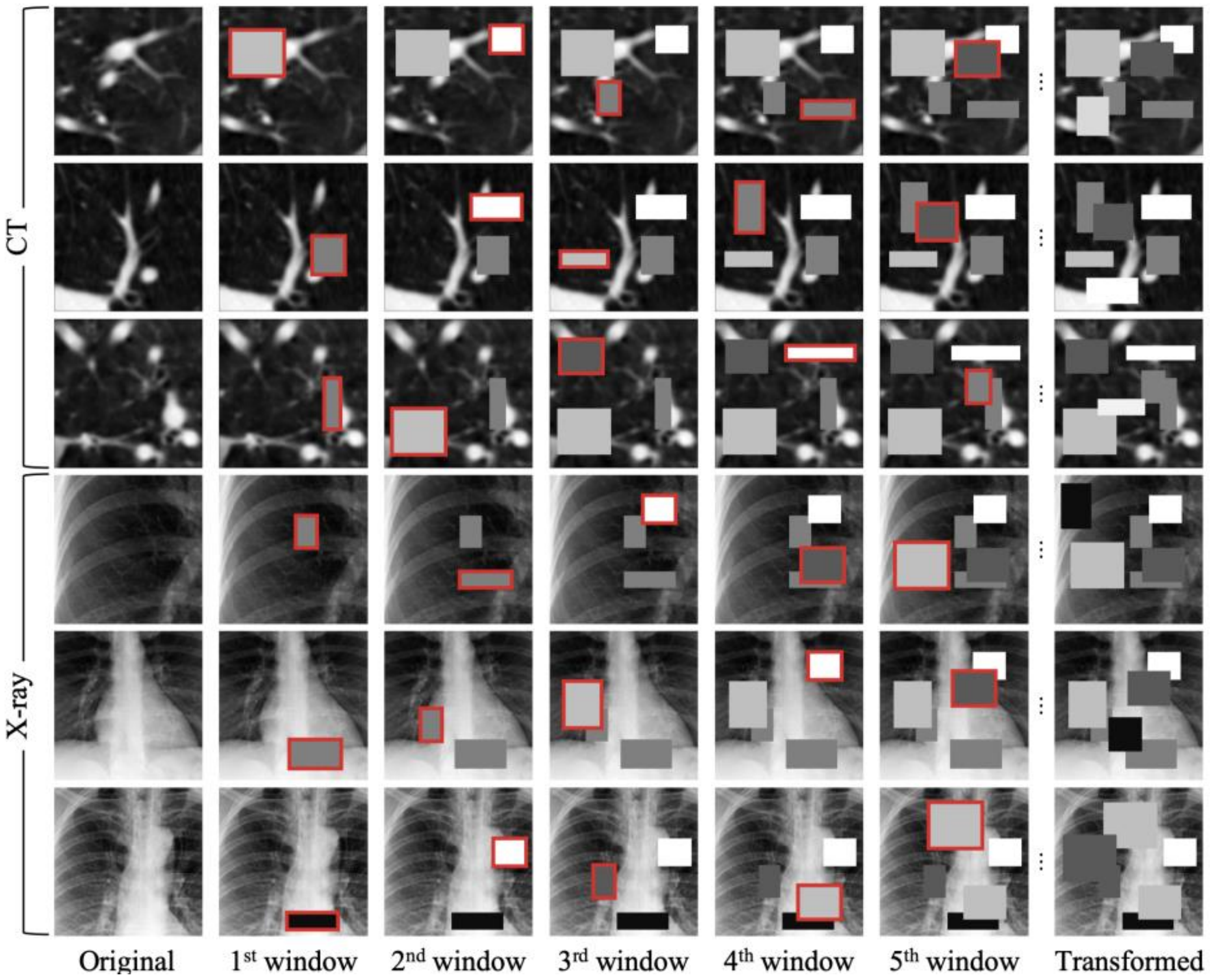
Out-painting & In-painting

1. Generate arbitrary number of windows of various sizes and aspect ratios.
2. Out-painting compels Models Genesis to learn global geometry and spatial layout of organs via extrapolating.
3. In-painting requires Models Genesis to appreciate local continuities of organs via interpolating.

Out-painting



In-painting


CT
X-ray
Original
1st window
2nd window
3rd window
4th window
5th window
Transformed

Model Genesis Properties

1. **Autodidactic**: requiring no manual labeling
2. **Eclectic**: learning from multiple perspectives (appearance, texture, context et)
3. **Scalable**: eliminating proxy-task-specific heads. By unifying all tasks as a single restoration task, any favorable transformation can be easily amended into the framework.
4. **Generic**: yielding diverse applications.
   a) Encoder for classification tasks
   b) Encoder-decoder for segmentation tasks

Experiments

1. All genesis models are self-supervised.
2. Genesis CT is pretrained from 534 CT scans in LIDC-IDRI
3. Genesis X-ray is pretrained from 77,074 X-rays in ChestXray8
4. Transferred to 7 downstream target tasks:
   a) 3D U-Net is used in five 3D applications
   b) U-Net architecture with ResNet-18 encoder is used in seven 2D applications

Table 1: Target tasks.

Code†	Object	Modality	Source	Description
NCC	Lung Nodule	CT	LUNA2016	Lung nodule false positive reduction
NCS	Lung Nodule	CT	LIDC-IDRI	Lung nodule segmentation
ECC	Pulmonary Embolism	CT	PE-CAD	Pulmonary embolism false positive reduction
LCS	Liver	CT	LiTS2017	Liver segmentation
DXC	Pulmonary Diseases	X-ray	ChestX-ray8	Eight pulmonary diseases classification
IUC	CIMT RoI	Ultrasound	UFL MCAEL	RoI, bulb, and background classification
BMS	Brain Tumor	MRI	BraTS2013	Brain tumor segmentation

† The first letter denotes the object of interest ("N" for lung nodule, "E" for pulmonary embolism, "L" for liver, etc); the second letter denotes the modality ("C" for CT, "X" for X-ray, "U" for Ultrasound, etc); the last letter denotes the task ("C" for classification, "S" for segmentation).

Table 2: Fine-tuning models from our Genesis Chest CT (3D) significantly outperforms learning from scratch in the five 3D target tasks ($p < 0.05$). The cells checked by ✗ denote the properties that are different between the proxy and target datasets. Our results show that our Genesis Chest CT generalizes across organs, diseases, datasets, and modalities. Footnotes show state-of-the-art performance for each target task.

Task	Metric	Disease	Organ	Dataset	Modality	Scratch (%)	Genesis (%)	p-value
NCC[1]	AUC					94.25±5.07	**98.20±0.51**	0.0180
NCS[2]	IoU					74.05±1.97	**77.62±0.64**	1.04e-4
ECC[3]	AUC	✗		✗		79.99±8.06	**88.04±1.40**	0.0058
LCS[4]	IoU	✗	✗	✗		74.60±4.57	**79.52±4.77**	0.0361
BMS[5]	IoU	✗	✗	✗	✗	90.16±0.41	**90.60±0.20**	0.0041

[1] LUNA winner holds an official score of 0.968 vs. 0.971 (ours)
[2] Wu *et al.* holds a Dice of 74.05% vs. 75.86%±0.90% (ours)
[3] Zhou *et al.* holds an AUC of 87.06% vs. 88.04%±1.40% (ours)
[4] LiTS winner w/ postprocessing (PP) holds a Dice of 96.60% vs. 91.13%±1.51% (ours w/o PP)
[5] BraTS winner w/ ensembling holds a Dice of 91.00% vs. 92.58%±0.30% (ours w/o ensembling)

Table 3: Comparison between our unified framework and each of the suggested self-supervised schemes on five 3D target tasks. The statistical analyses is conducted between the top-2 models in each column highlighted in red. While there is no clear winner, our unified framework is more robust across all target tasks, yielding either the best result or comparable performance to the best model ($p > 0.05$).

Approach	NCC (%)	NCS (%)	ECC (%)	LCS (%)	BMS (%)
Scratch	94.25±5.07	74.05±1.97	79.99±8.06	74.60±4.57	90.16±0.41
Distortion (ours)	96.46±1.03	77.08±0.68	**88.04±1.40**	79.08±4.26	**90.60±0.20**
Painting (ours)	**98.20±0.51**	77.02±0.58	87.18±2.72	78.62±4.05	90.46±0.21
Unified (ours)	97.90±0.57	**77.62±0.64**	87.20±2.87	**79.52±4.77**	90.59±0.21
p-value	0.0848	0.0520	0.2102	0.4249	0.4276

Table 4: Comparison between 3D solutions and 2D slice-based solutions on three 3D target tasks. Training 3D models from scratch does not necessarily outperform the 2D counterparts (see NCC). However, training the same 3D models from Genesis Check CT outperforms ($p < 0.05$) all 2D solutions, demonstrating the effectiveness of Genesis Chest CT in unlocking the power of 3D models.

Task	2D (%)			3D (%)			p-value[†]
	Scratch	ImageNet	Genesis	Scratch	ImageNet	Genesis	
NCC	96.03±0.86	97.79±0.71	97.45±0.61	94.25±5.07	N/A	**98.20±0.51**	0.0213
NCS	70.48±1.07	72.39±0.77	72.20±0.67	74.05±1.97	N/A	**77.62±0.64**	<1e-8
ECC	71.27±4.64	78.61±3.73	78.58±3.67	79.99±8.06	N/A	**88.04±1.40**	5.50e-4

[†]These p-values are calculated between our Models Genesis vs. the fine-tuning from ImageNet, which always offers the best performance (highlighted in red) for all three tasks in 2D.

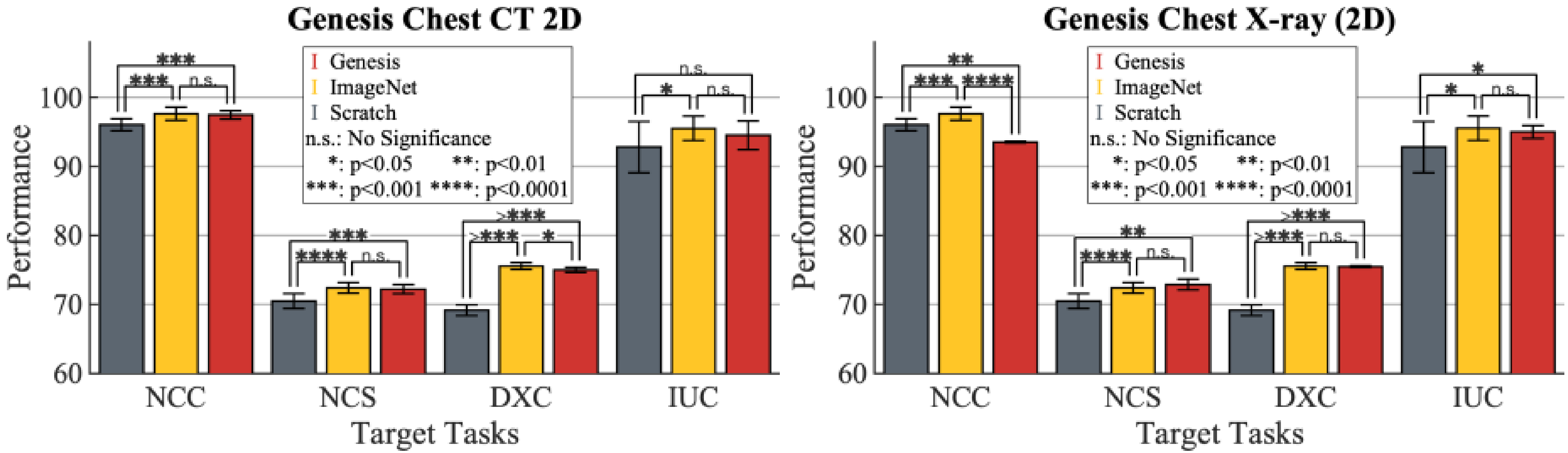
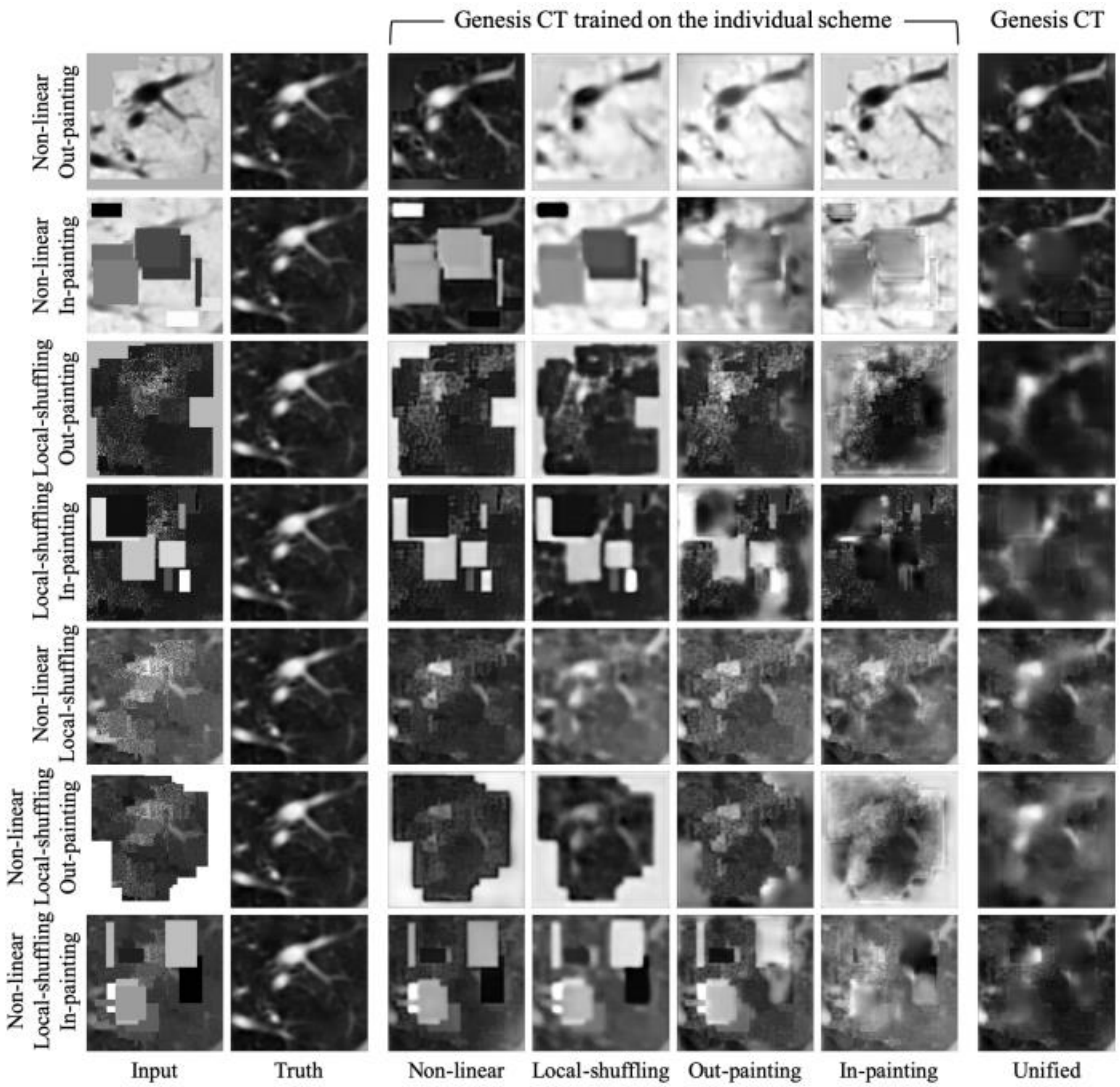


Fig. 2: Comparison of 2D solutions on four 2D target tasks. To investigate the same- and cross-domain transferability of Models Genesis, we have trained Genesis Chest CT 2D using 2D axial slices from LUNA dataset (left panel), and Genesis Chest X-ray (2D) trained using radiographs from ChestX-ray8 dataset (right panel). In same-domain target tasks (NCC and NCS in the left panel and DXC in the right panel), Models Genesis 2D outperform training from scratch and offer equivalent performance to fine-tuning from ImageNet. While in cross-domain target tasks (DXC and IUC in the left panel; NCS and IUC in the right panel), Models Genesis 2D also produce fairly robust performance.

Conclusions

1. Models Genesis
   a) Outperform 3D models trained from scratch
   b) Consistently top any 2D approaches
   c) (2D) offer equivalent performances to supervised pretrained models
2. Future work:
   a) Modality-oriented models, such as Genesis MRI and Genesis Ultrasound
   b) Organ-oriented models, such as Genesis Brain and Genesis Heart


Genesis CT trained on the individual scheme
Genesis CT
Non-linear Out-painting
Non-linear In-painting
Local-shuffling Out-painting
Local-shuffling In-painting
Non-linear Local-shuffling
Non-linear Local-shuffling Out-painting
Non-linear Local-shuffling In-painting
Input
Truth
Non-linear
Local-shuffling
Out-painting
In-painting
Unified