(TUE-PM-181)

PosterLayout: A New Benchmark and Approach for Content-aware Visual-Textual Presentation Layout

HsiaoYuan Hsu[1,2], Xiangteng He[1,2], Yuxin Peng[1,2], Hao Kong[3] and Qing Zhang[3]

[1]Wangxuan Institute of Computer Technology, Peking University
[2]National Key Laboratory for Multimedia Information Processing, Peking University
[3]Meituan

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

(TUE-PM-181)

PosterLayout: A New Benchmark and Approach for Content-aware Visual-Textual Presentation Layout

HsiaoYuan Hsu[1,2], Xiangteng He[1,2], Yuxin Peng[1,2], Hao Kong[3] and Qing Zhang[3]

[1]Wangxuan Institute of Computer Technology, Peking University

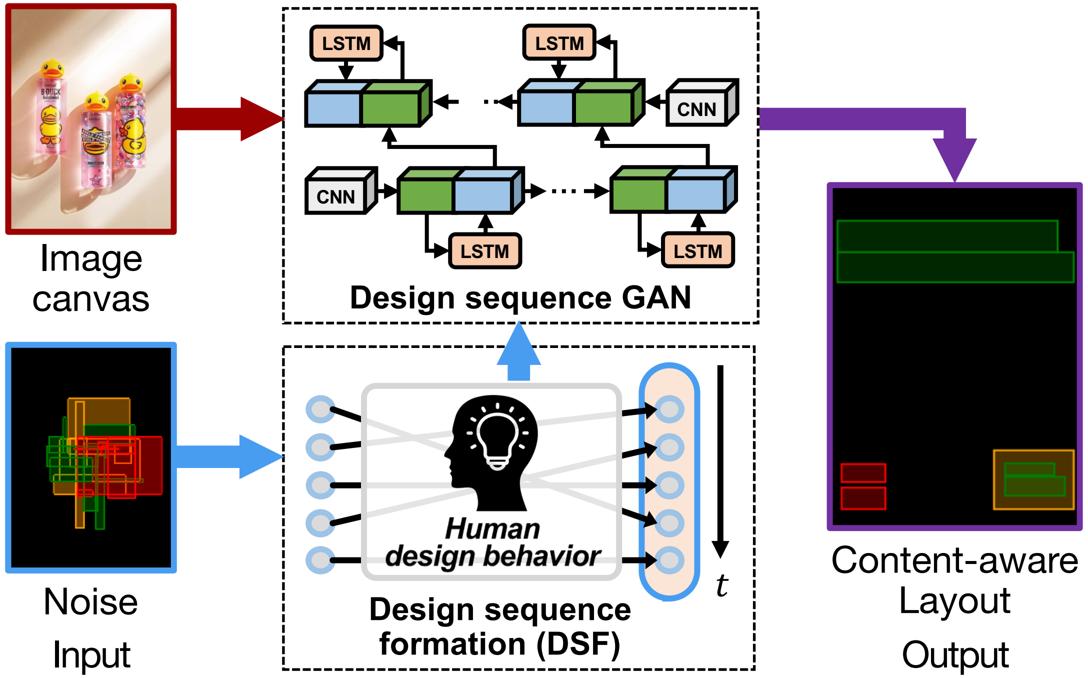
[2]National Key Laboratory for Multimedia Information Processing, Peking University [3]Meituan

Project page

Go!

Training set:
Poster-layout pair × 9,974

Test set:
Image canvas × 905



PKU PosterLayout Dataset

Design Sequence GAN (DS-GAN)

Outline

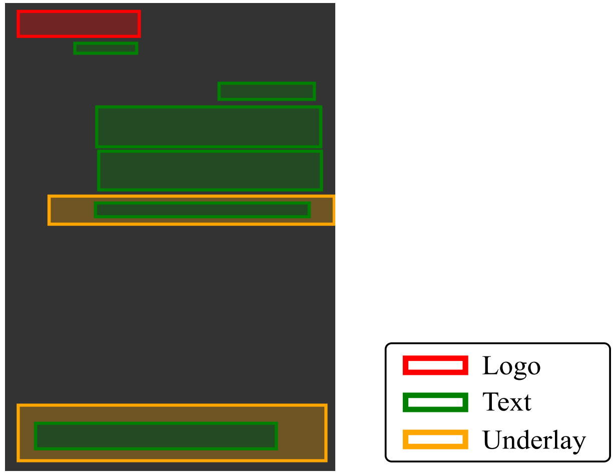
- Introduction
- A New Benchmark: PKU PosterLayout
- A New Approach: Design Sequence GAN (DS-GAN)
- Experiments
- Conclusion

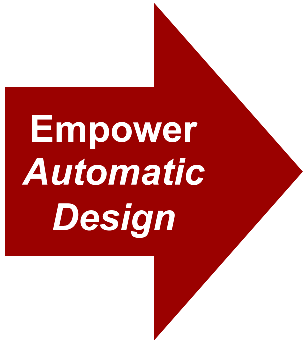
Background & Application Scenario

- Content-aware Visual-Textual Presentation Layout
  - Given an **image canvas**, arrange spatial space for *<u>informative or decorative elements</u>*, such as text, logo, and underlay
  - Useful in template-free poster designs

Task input:
Content-rich Canvas

Task output:
Content-aware Layout

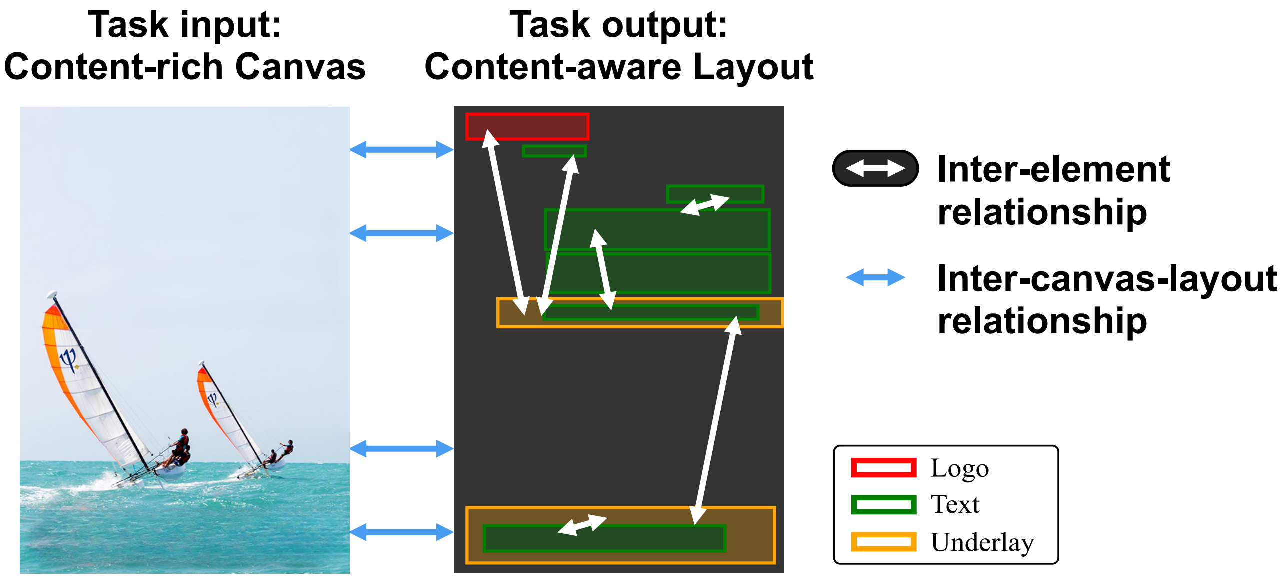




+ desired contents

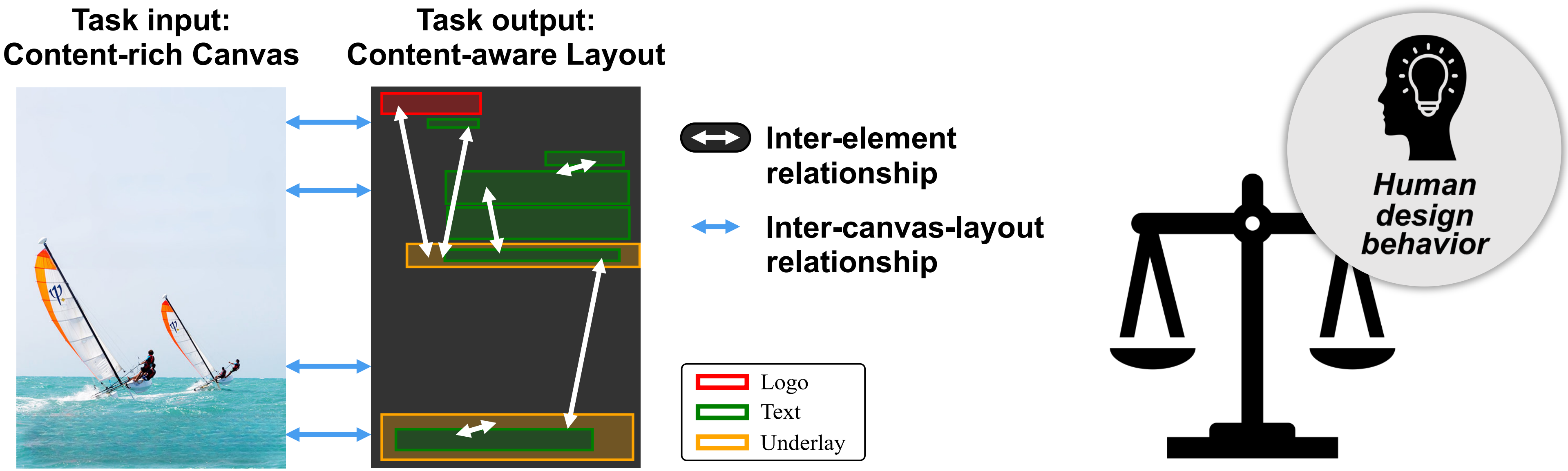
Challenges & Motivations (1/2)

- Complex **inter-element** and **inter-canvas-layout relationships** modeling
  - Considering the two relationships in a balanced manner is critical
  - Human design behavior can provide a naturally balanced heuristics
- Lack of a **public benchmark** dedicated to this novel task



Challenges & Motivations (2/2)

- Complex **inter-element** and **inter-canvas-layout relationships** modeling
  - Considering the two relationships in a balanced manner is critical
  - Human design behavior can provide a naturally balanced heuristics
- Lack of a **public benchmark** dedicated to this novel task



Outline

- Introduction
- **A New Benchmark: PKU PosterLayout**
- A New Approach: Design Sequence GAN (DS-GAN)
- Experiments
- Conclusion

A New Benchmark: PKU PosterLayout

Training set:
Poster-layout pair × 9,974

Test set:
Image canvas × 905

Specialties of PKU PosterLayout (1/3)

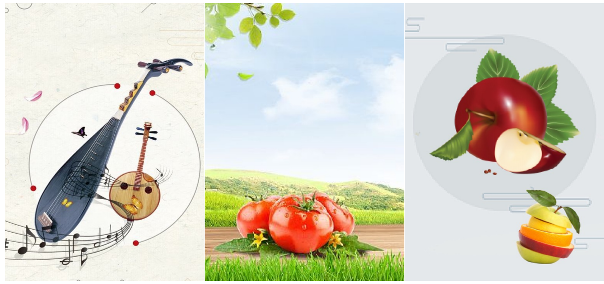
i. Domain diversity

- Data were collected from multiple sources, varying in **domain**, **quality**, and **resolution**
  - Shifts in distributions can make the dataset more **general**

Image sources
(1) *https://www.taobao.com/* [1] (2) *https://unsplash.com/* (3) *https://www.freepik.com/* (4) *https://pixabay.com/* (5) *https://pngimg.com/* (6) *https://www.stickpng.com/*

Natural images

E-commerce product images



Blended images

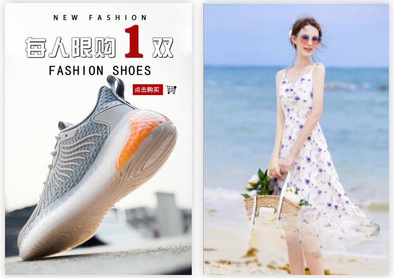
Specialties of PKU PosterLayout (2/3)

ii. Content diversity

- Objects in images are broadly distributed in **9 coarse-grained categories** covering most e-commerce products

Food / drinks

Fresh produce



Clothing

Toys / instruments

Groceries

Sports / transportation

Cosmetics / accessories

Electronics / office supplies

Appliances / decor

Specialties of PKU PosterLayout (3/3)

iii. Layout complexity and variety

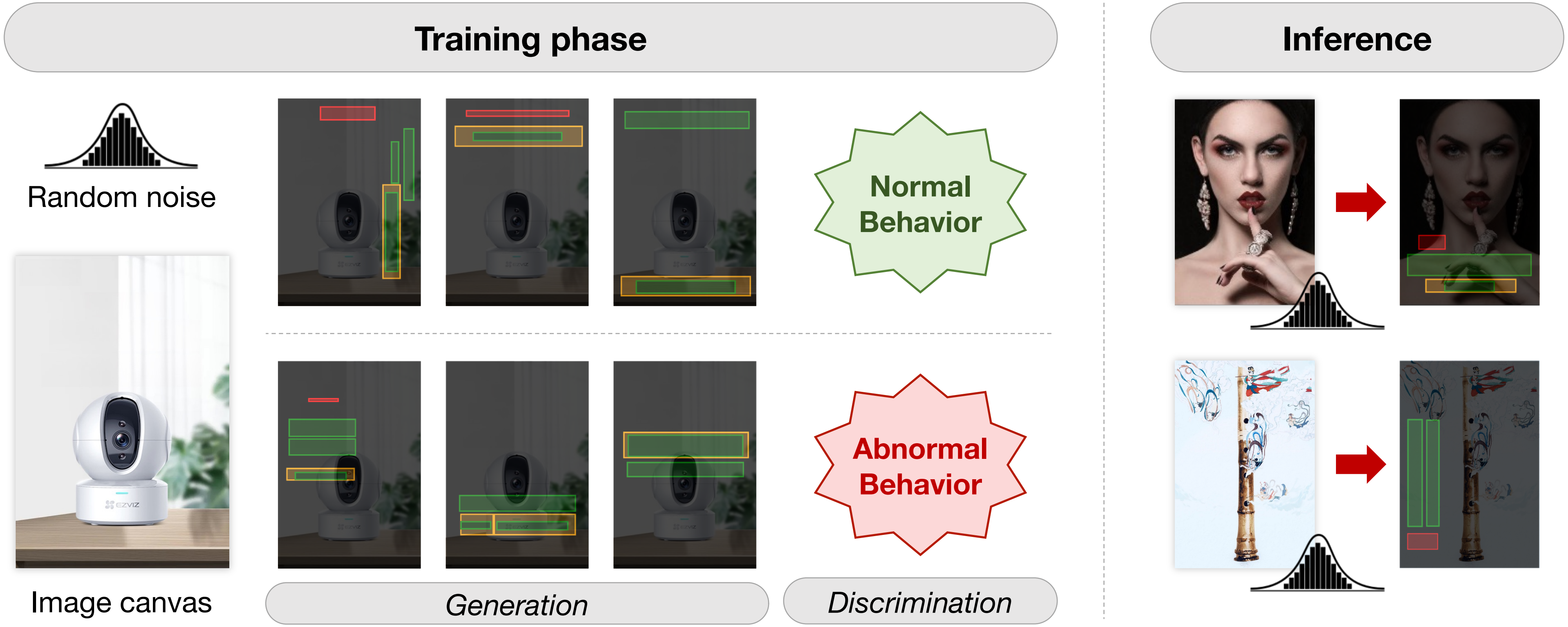
- It is the ***first*** public dataset containing **complex layouts** with >10 elements
  - Providing more difficulties in modeling the inter-element relationship
  - Capable of supporting ***extended tasks*** requiring complex layouts

Outline

- Introduction
- A New Benchmark: PKU PosterLayout
- **A New Approach: Design Sequence GAN (DS-GAN)**
- Experiments
- Conclusion

A New Approach: Design Sequence GAN (DS-GAN)

- Abstract design behavior into the order in which the designer places elements on the canvas, named as **design sequence**

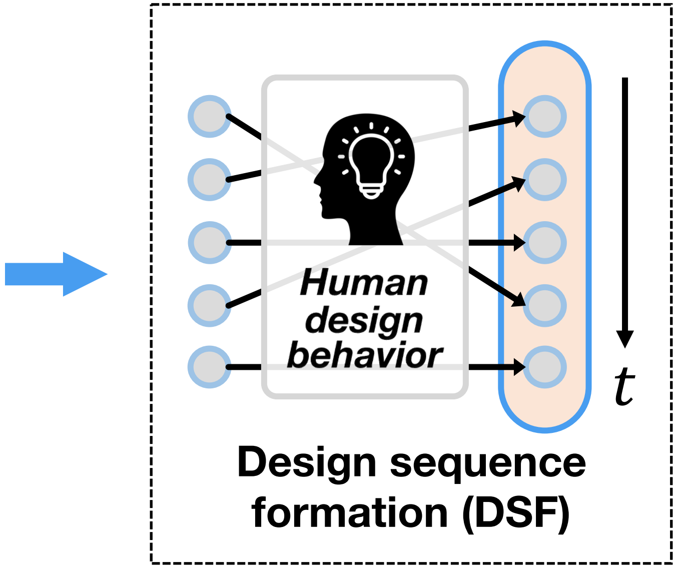


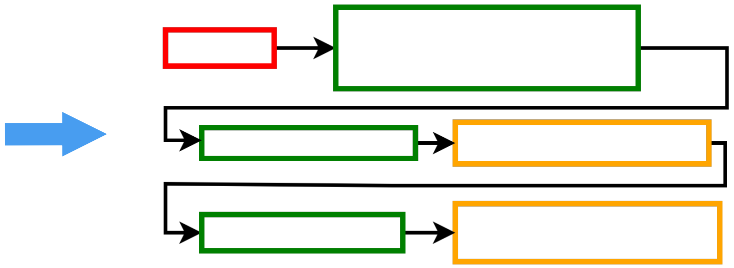
Design Sequence GAN (1/3)

i. Design sequence formation (DSF)

- Inspired by **Human design behavior**
- Converting plain layout data into temporal design sequences
  - Considering **(1) category**, **(2) area**, and **(3) grouping** of elements

Plain layout data





Design sequence

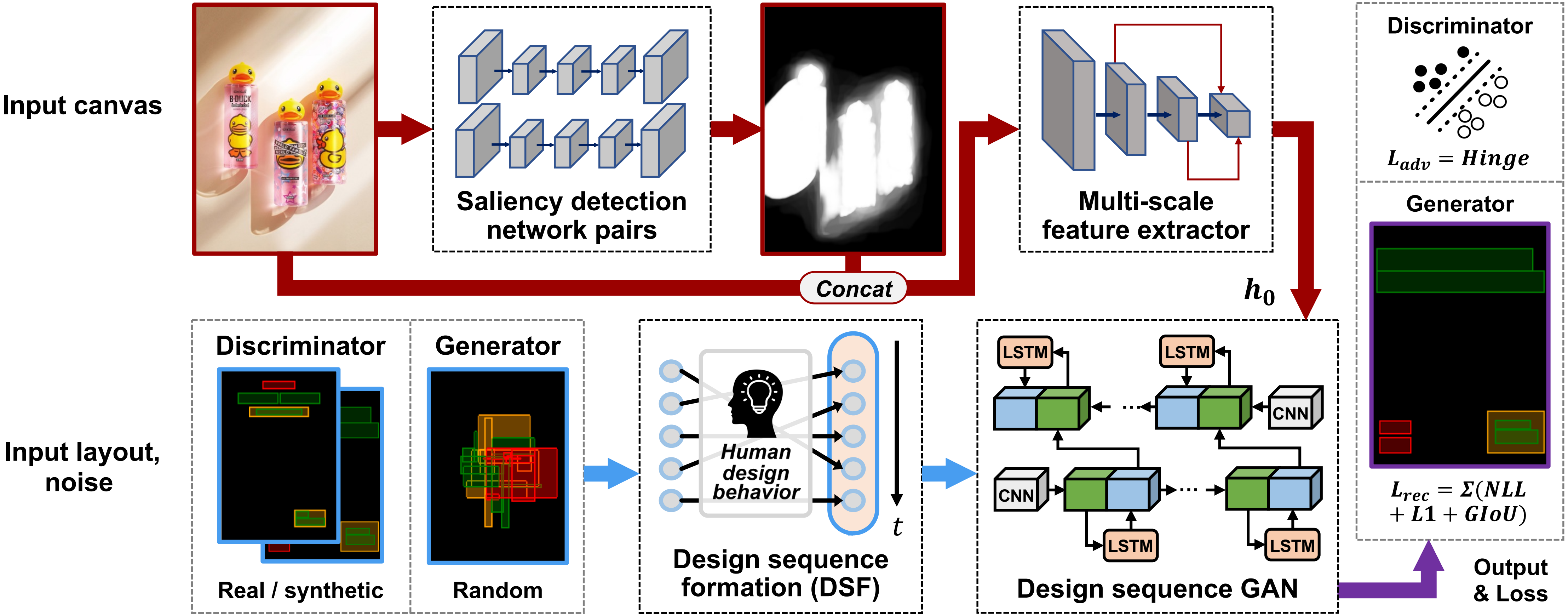
Design Sequence GAN (2/3)

ii. Design sequence GAN (DS-GAN)

- Implemented by CNN-LSTM models, triggered by visual features of the canvas to generate image content-aware layouts
- Acting like a human who **first observes the image and then starts the design**

Design Sequence GAN (3/3)

ii. Design sequence GAN (DS-GAN)



Outline

- Introduction
- A New Benchmark: PKU PosterLayout
- A New Approach: Design Sequence GAN (DS-GAN)
- **Experiments**
- Conclusion

Comparisons with State-of-the-art Methods

- Compare our DS-GAN with SOTA methods on **PKU PosterLayout** and report metrics, including:
  - ***Graphic*** metrics: evaluating inter-element relationship

    {Validity of size, Overlay, Alignment, Underlay effectiveness (loose, strict)}
  - ***Content-aware*** metrics: evaluating inter-canvas-layout relationship

    {Utilization rate of non-salient region, Occlusion, Readability}

	Target	*Val* ↑	*Ove* ↓	*Ali* ↓	*Und_l* ↑	*Und_s* ↑	*Uti* ↑	*Occ* ↓	*Rea* ↓
SmartText [2]	T	-	-	-	-	-	0.0849	**0.0912**	**0.1528**
CGL-GAN [3]	V-T	0.7066	0.0605	0.0062	**0.8624**	0.4043	0.2257	0.1546	0.1715
DS-GAN (Ours)	V-T	**0.8788**	**0.0220**	**0.0046**	0.8315	**0.4320**	**0.2541**	0.2088	0.1874

Almost **dominate** graphic metrics

Get a **good trade-off** between two aspects of metrics

Ablation Study (1/2)

- Gain insight into the effects of
  - CNN-LSTM models: remaining only the last fully connected layers

	Val ↑	*Ove* ↓	*Ali* ↓	*Und_l* ↑	*Und_s* ↑	*Uti* ↑	*Occ* ↓	*Rea* ↓
Without CNN-LSTM	0.6765	0.0888	0.0112	0.0106	0.0000	0.2155	0.2804	0.2015
With CNN-LSTM (DS-GAN)	**0.8788**	**0.0220**	**0.0046**	**0.8315**	**0.4320**	**0.2541**	**0.2088**	**0.1874**

Removing the **behavior pattern model** destroys the methodology

Ablation Study (2/2)

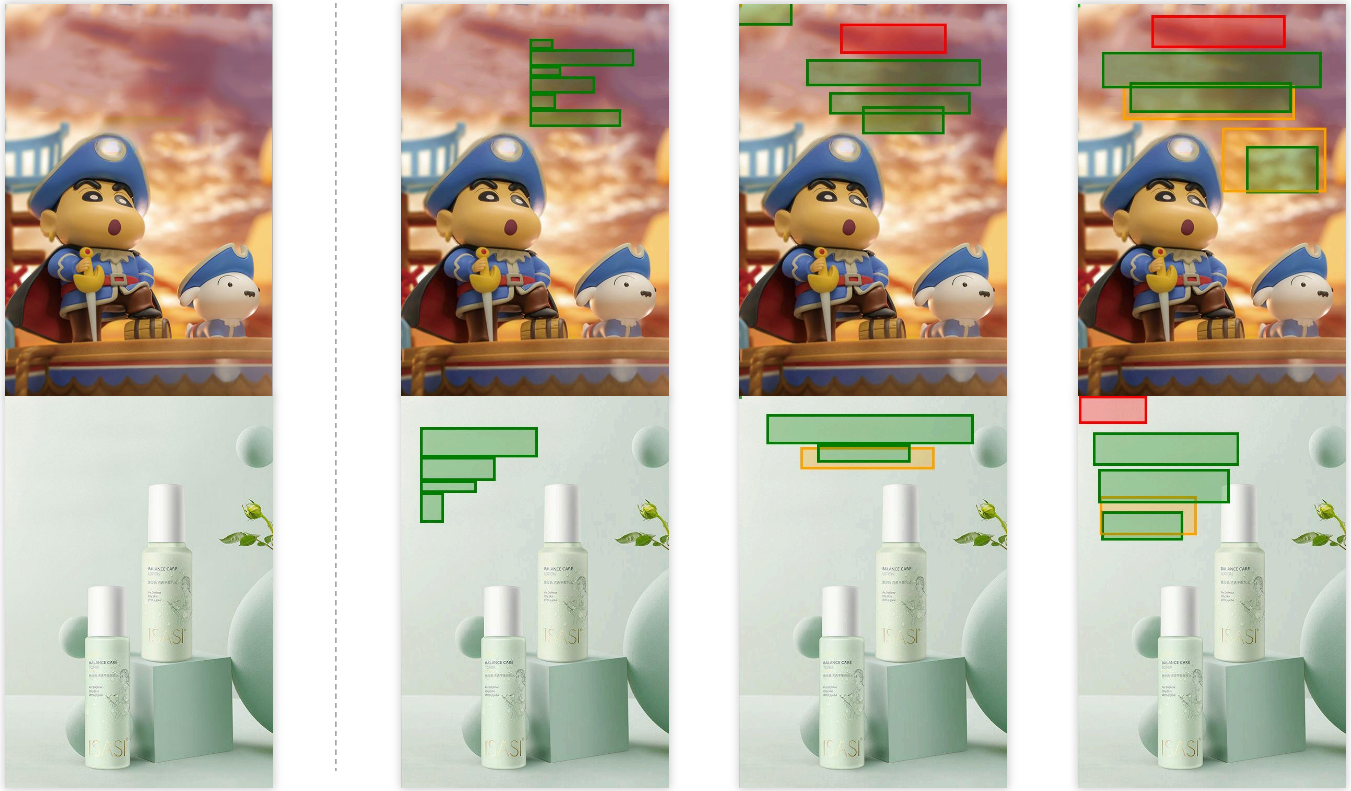
- Gain insight into the effects of
  - **CNN-LSTM models**: remaining only the last fully connected layers
  - **DSF**: limiting the maximum sequence length and adopting different formation strategies

	Val ↑	*Ove* ↓	*Ali* ↓	*Und_l* ↑	*Und_s* ↑	*Uti* ↑	*Occ* ↓	*Rea* ↓	*AE* ↓
Random	**1.000** (+0.1454)	0.0881 (+0.0666)	0.0062 (+0.0007)	0.7417 (-0.1380)	0.3243 (-0.1499)	0.2240 (-0.0328)	0.2475 (+0.0361)	**0.1909** (+0.0035)	0.5730
Geometric	0.9667 (+0.1215)	**0.0261** (+0.0026)	0.005 (+0.0004)	0.7849 (-0.0824)	0.4433 (-0.0757)	0.2439 (-0.0170)	0.2482 (+0.0438)	0.1937 (+0.0052)	0.3486
DSF-based (DS-GAN-*8*)	0.9572 (+0.0784)	0.0362 (+0.0142)	**0.0043** (-0.0003)	**0.8850** (+0.0535)	**0.5824** (+0.1504)	**0.2526** (-0.0015)	**0.2341** (+0.0253)	0.1910 (+0.0036)	**0.3272**
	More effective								Stabler

Visualized Results (1/3)

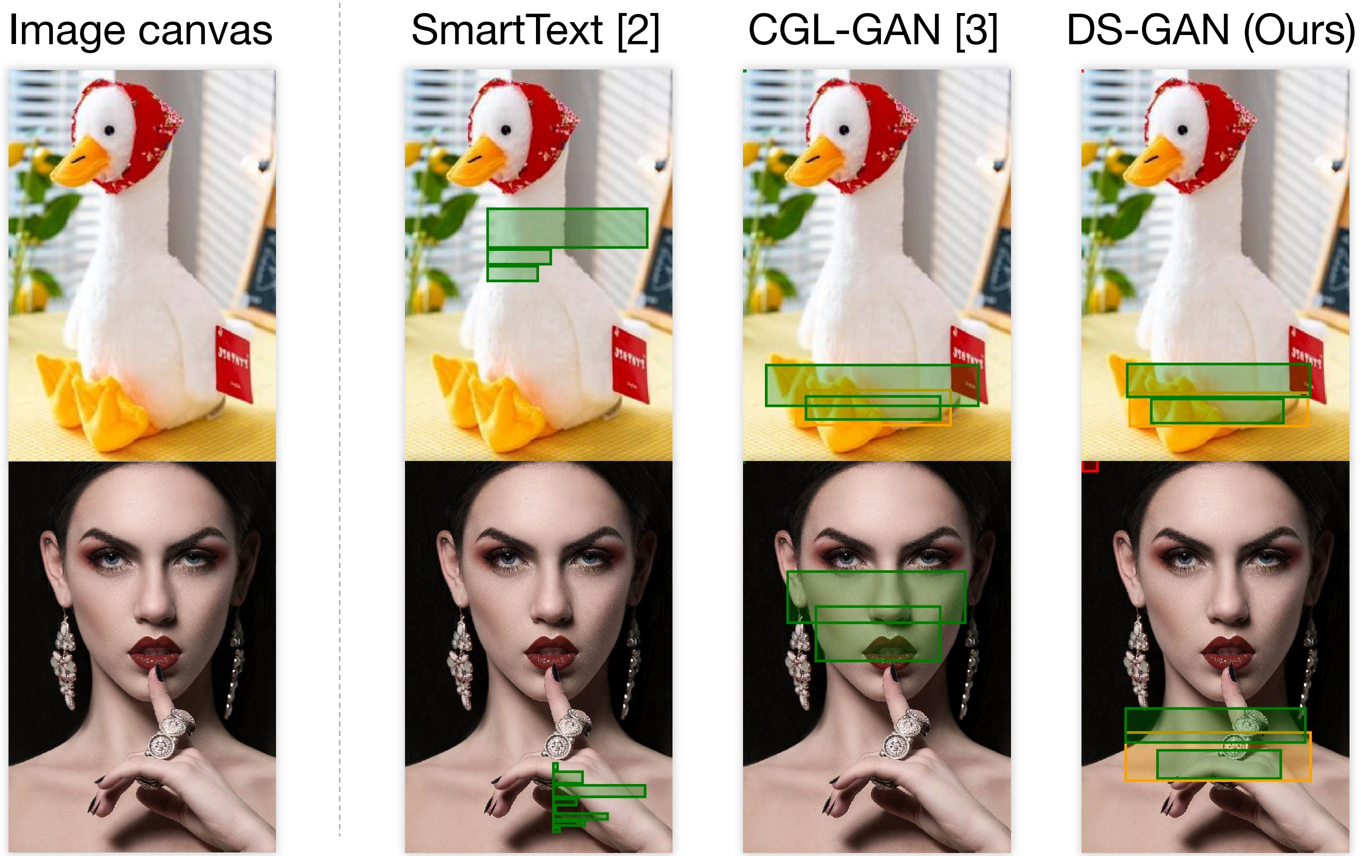
- Our DS-GAN generates more appealing layouts for diverse canvases
  - Actively **utilize** all suitable spaces, retaining some *visually natural* **occlusion**

Image canvas | SmartText [2] | CGL-GAN [3] | DS-GAN (Ours)



Visualized Results (2/3)

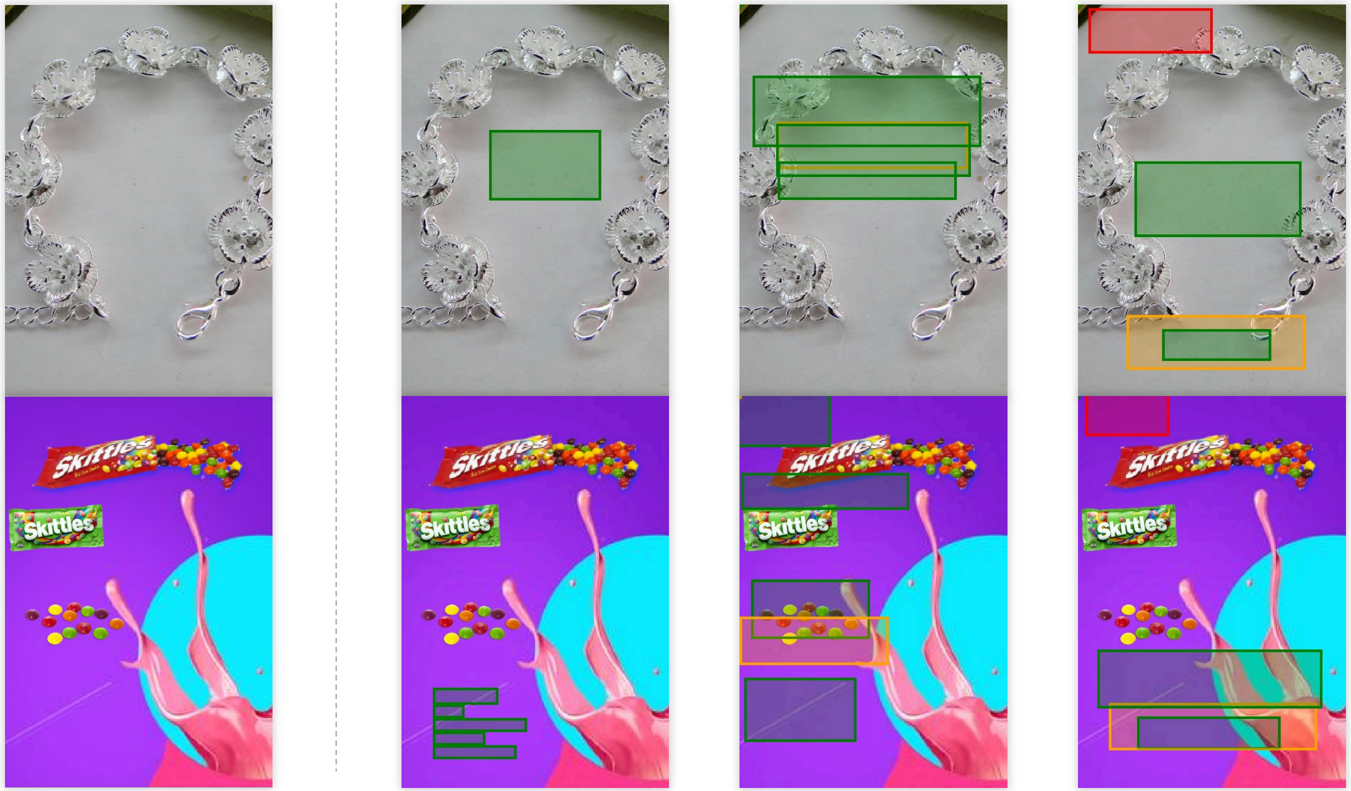
- Our DS-GAN generates more appealing layouts for diverse canvases
  - Avoid *unpleasant* **overlay**, **non-alignment**, or **occlusion**



Visualized Results (3/3)

- Our DS-GAN generates more appealing layouts for diverse canvases
  - Capable of handling canvases with **special-shaped, complex objects**

Image canvas | SmartText [2] | CGL-GAN [3] | DS-GAN (Ours)



Outline

- Introduction
- A New Benchmark: PKU PosterLayout
- A New Approach: Design Sequence GAN (DS-GAN)
- Experiments
- **Conclusion**

Conclusion

- This paper devoted to content-aware visual-textual presentation layouts by
  - Construct a new benchmark, ***PKU PosterLayout***
  - Propose a new generative approach, ***DS-GAN***, inspired by human behavior
    - Composed of DSF and CNN-LSTM-based GAN, both of which are critical
- Several experiments were conducted and verified ***PKU PosterLayout***'s usefulness and ***DS-GAN***'s effectiveness
- The dataset and code are open-sourced *(visit the project page!)*, hopefully encouraging further research

Reference

[1] Gangwei Jiang, Shiyao Wang, Tiezheng Ge, Yuning Jiang, Ying Wei, and Defu Lian. Self-supervised text erasing with controllable image synthesis. In Proceedings of the ACM International Conference on Multimedia (ACM MM), pages 1973–1983, 2022.

[2] Chenhui Li, Peiying Zhang, and Changbo Wang. Harmonious textual layout generation over natural images via deep aesthetics learning. IEEE Transactions on Multimedia (TMM), 2021.

[3] Min Zhou, Chenchen Xu, Ye Ma, Tiezheng Ge, Yuning Jiang, and Weiwei Xu. Composition-aware graphic layout GAN for visual-textual presentation designs. In Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI), pages 4995–5001, 2022.

(TUE-PM-181)

PosterLayout: A New Benchmark and Approach for Content-aware Visual-Textual Presentation Layout

HsiaoYuan Hsu[1,2], Xiangteng He[1,2], Yuxin Peng[1,2], Hao Kong[3] and Qing Zhang[3]

[1]Wangxuan Institute of Computer Technology, Peking University [2]National Key Laboratory for Multimedia Information Processing, Peking University [3]Meituan

Thank you!

& Feel free to contact us!!

[Lab homepage] [GitHub page] [WeChat page]

Project page

Go!