KAN-Attn GAN: Map Generation with Kolmogorov-Arnold Networks and Attention-Based Queries Selection

Arpan Mahara*, Naphtali D. Rishe†, Wenjia Wang‡, Seyed Masoud Sadjadi§
Knight Foundation School of Computing and Information Sciences
Florida International University
Miami, FL 33199
*amaha038@fiu.edu, †rishen@cs.fiu.edu, ‡wwang048@fiu.edu, §sadjadi@cs.fiu.edu

2024 International Conference on Machine Learning and Applications (ICMLA) | 979-8-3503-7488-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICMLA61862.2024.00247

***Abstract*—Dynamic urban development and natural disasters often lead to discrepancies between maps and satellite images, posing challenges for autonomous navigation and smart city planning. Automatic map generation from satellite imagery is crucial for addressing these discrepancies and ensuring accurate geospatial data. While Generative Adversarial Networks (GANs) favored with contrastive learning have been used for this task, limitations such as randomness in feature selection result in suboptimal performance, and MLP-based projection heads lack the expressiveness needed for effective feature generation. To address these limitations, we propose a novel model, KAN-Attn, which incorporates an attention-based query selection mechanism for selecting relevant features in contrastive learning. Also, KAN-Attn is the first application of Kolmogorov-Arnold Network (KAN) to improve feature generation with enhanced expressiveness in map generation. Our innovations allow KAN-Attn to achieve optimal performance in map generation, with experiments on publicly available datasets demonstrating state-of-the-art results based on relevant metrics on perceptual quality.**

***Index Terms*—Kolmogorov-Arnold Network (KAN), Generative Adversarial Networks (GANs), Map Generation, Contrastive Learning, Query Selection, Attention Mechanism**

I. Introduction

Satellite images capture a variety of underlying objects such as roads, green areas, water bodies, and buildings. Transforming satellite images into human-readable formats, such as maps, is crucial for improving accessibility and supporting a wide range of applications. Typically, these transformed map images are considered vector images, offering a clearer representation of key features. Historically, the conversion from raster to vector images has been a manual, labor-intensive process, prone to human error.

To automate this process, generative AI techniques have gained considerable attention in recent years. The focus has shifted from manually constructing maps to automatically generating them using generative models [15], [3]. However, the scarcity of paired satellite and map images presents a major challenge, making supervised learning approaches impractical. Generative Adversarial Networks (GANs) [8] have emerged as a leading solution for this challenge, particularly for unsupervised map generation. Several studies [7], [27], [3], [28], and [21] have demonstrated the success of GANs in generating map images from satellite data. However, the inherent complexity of satellite imagery introduces significant limitations, such as obstacles like shadows, trees, or buildings obscuring key features.

Fig. 1: **Example Satellite and Corresponding Map Images.** Green arrow shows satellite to map translation from [15], and red arrow depicts styled maps to simple maps from [16].

While GANs enhanced with contrastive learning [26], [21] outperform traditional GANs in map generation, they still suffer from key limitations, including the randomness in feature selection and the use of MLP projection heads [13], which lack sufficient expressiveness. Furthermore, although MLPs are prevalent in various domains for system development, they may face security challenges [24], [11], [32], [10].

To address these challenges, we explore the potential of Kolmogorov-Arnold Networks (KANs) [20], a recent development in deep learning proposed as a powerful alternative to the conventional MLP. KANs have demonstrated strong expressiveness and achieved success in satellite image classification tasks [5]. Despite this, their application in generative tasks, particularly in map generation, remains unexplored.

Recognizing both the limitations of GANs and the potential of KANs, we hypothesize that KAN can significantly improve map generation from satellite images. Our study presents the following key innovations:

- First application of KAN as a projection head in map generation, enhancing feature generation.

1946-0759/24/$31.00 ©2024 IEEE
DOI 10.1109/ICMLA61862.2024.00247

- Incorporation of an attention-based query selection mechanism to identify and select important features, boosting the contrastive learning process.
- Development of a novel model, KAN-Attn, achieving state-of-the-art performance in map generation, as evidenced by FID and KID metrics.

II. Related Work

Several studies have explored the automatic generation of maps from satellite images using generative models. The success of map construction using generative mechanisms was advanced with CycleGAN [34], which enabled image translation between two domains without requiring paired datasets. GeoGAN [7] extended adversarial learning by integrating reconstruction and style transfer losses, successfully generating standard layer maps. Despite its advances, GeoGAN struggled with the coarse resolution of satellite images, limiting the quality of the generated maps. Similarly, SMAPGAN [3] incorporated a semi-supervised strategy, allowing the model to be pretrained on unpaired satellite and map images and fine-tuned with a small set of paired images. This model introduced gradient-based L1 loss and gradient structure loss, preserving topological coherence and edge details.

Furthermore, MapGen-GAN [28] built on CycleGAN's cycle-consistency principle, adding circularity-consistency and geometrical-consistency constraints to improve semantic and geometric precision. However, its dependence on geometric transformations limits flexibility in diverse tasks. CscGAN [19] employed multi-scale generators and discriminators to generate tile maps at different resolutions, but its reliance on scale-specific generators restricts adaptability in tasks needing uniform resolution. MapCUT [21] introduced contrastive learning into the map generation process, replacing CycleGAN's cyclic nature with a unidirectional approach. Its use of patch-based contrastive learning (PatchNCE) improved alignment between satellite images and maps, yet the two-layer MLP limited feature expressiveness. Subsequently, HPix [31] proposed a hierarchical GAN architecture for generating vector maps, using a global GAN for overall layout and a local GAN for refining details. While effective, its reliance on two distinct generators posed challenges for scalability and generalization

Despite these advancements, current methods still struggle with feature expressiveness, accuracy and scalability. To address these gaps, we propose the KAN-Attn model, which leverages KAN and attention-based query selection for more accurate and robust map generation.

III. Method

This section begins by outlining the general use of contrastive learning in GANs, followed by the application of query-selection attention and the integration of KAN to introduce the KAN-Attn model.

A. Map generation setup

Consider a scenario where we have satellite images, denoted as X (raster images), and corresponding map images, denoted as Y (vector images), collected from specific geographic locations L. These satellite images, X, contain multiple layers such as roads, buildings, and water bodies, and may change over time due to factors like natural disasters. Our goal is to generate a map, y, that accurately reflects these altered layers in the satellite image $x \in X$ at location $l \in L$. The objective is to train a generator G, utilizing GANs and contrastive learning, to effectively capture and generalize the features of these layers from domain X and translate them into accurate map representations in domain Y.

The main challenge is the lack of paired satellite and map images, as well as labeled object layers, making this an unsupervised learning task. To address this, we aim to train G to learn these features in a self-supervised manner through one-directional training, drawing on both GAN and contrastive learning principles.

B. Generator of GAN with Contrastive Learning

The generative model consists of two neural networks: a generator and a discriminator, forming the core of the GAN architecture. We adopt a one-directional GAN, implemented with a single generator and discriminator, following the architectures presented in [26] and [21]. For the generator, we utilize an encoder-decoder design based on ResNet as depicted in Fig. 2. The generator utilizes a ResNet-based encoder-decoder structure, with the encoder extracting enriched features from the satellite images, which the decoder translates into detailed map images, capturing the key object layers.

As mentioned earlier, our work is unsupervised, meaning there are no specific labels or paired images. Since the model is unsupervised, we work without labeled data or paired images. The challenge lies in generating map images that accurately reflect the underlying object layers in the satellite images. These object layers refer to physical entities such as roads, buildings, and water bodies, rather than neural network layers. To achieve this, we employ patch-level contrastive learning, maximizing the shared information between aligned patches in satellite and map images, minimizing the distance between similar patches in the embedding space. This approach is adapted from [2], [26], and [21].

We utilize the technique of future prediction as proposed in [25], where representation learning can help to predict positive features for the given queries and maximize mutual information between corresponding patches. Maximizing mutual information allows us to generate the corresponding object layer from the satellite image. The NT-Xent loss (normalized temperature-scaled cross-entropy loss) is used to correctly select the positive feature among a pool of $N-1$ negatives, further maximizing mutual information. The loss is formulated as:

$$L(q, v^+, v^-) = -\log \frac{\exp(\text{sim}(q, v^+)/\tau)}{\exp(\text{sim}(q, v^+)/\tau) + \sum_{i=1}^{N} \exp(\text{sim}(q, v_i^-)/\tau)} \quad (1)$$

where $\text{sim}(q, v^+)$ represents the cosine similarity between q and v^+, calculated as $\text{sim}(q, v^+) = \frac{q^\top v^+}{\|q\|\|v^+\|}$. The parameter τ represents a temperature factor that scales the distances between vectors. This procedure helps to identify the right

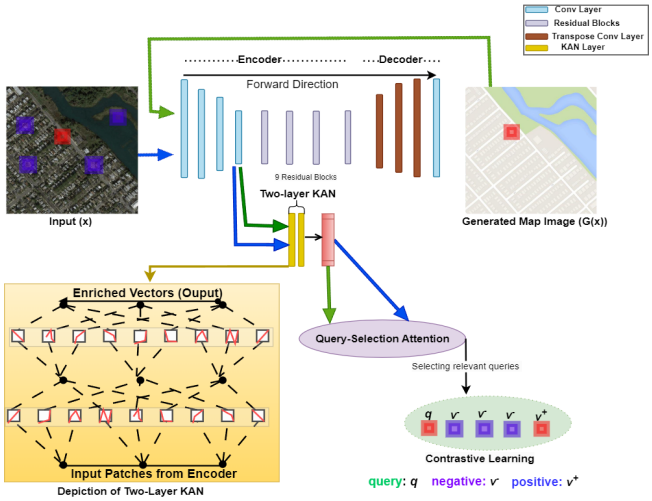

Fig. 2: **Generator Architecture for KAN-Attn.** The encoder-decoder generator integrates a two-layer KAN for feature enrichment and an attention mechanism for query selection to enhance contrastive learning. Satellite and map images are reproduced from [15] under the Creative Commons Attribution-NonCommercial 4.0 International License.

positive feature for a given query point, but the selection of the query point is random.

As the generator is a neural network model, the encoder processes images through multiple layers. When performing contrastive learning for mutual information maximization, [2] demonstrated that applying contrastive learning directly to the features extracted is less effective. Instead, an external projection head, denoted as H, is used to process these features. Typically, H is a two-layer MLP, which enriches feature understanding and improves results during inference. This concept of enriching features before contrastive learning has been explored in various studies in generative AI [26], [9], and [21].

While these previous methods can successfully generate maps from satellite images, we identify two key limitations in the feature learning process: 1. Contrastive learning randomly selects a query point, which can lead to suboptimal results if the query originates from a region without meaningful objects (e.g., empty fields or oceans). In such cases, the contrastive learning process may attempt to maximize information for irrelevant queries, negatively impacting the quality of the generated maps. 2. The two-layer MLP used as the projection head H, though effective, lacks the expressiveness needed for more robust feature generation. To address these limitations, we propose two improvements. First, we apply an attention mechanism to select relevant anchor points, denoted as R, by measuring the relationship between a query and its surrounding keys. This addresses the issue of random query selection. Second, we replace the two-layer MLP with a two-layer KAN (Kolmogorov-Arnold Network) to generate richer features, improving the feature extraction process. These innovations form the basis of our proposed KAN-Attn model.

C. KAN-Attn model

As demonstrated in [22], the KAN has shown its potential in improving generative AI tasks. We extend this approach to the task of map generation by replacing the two-layer MLP with a two-layer KAN (see Fig 2, where the section showing the two-layer KAN is highlighted), which provides more expressive feature generation for the decoder. For a detailed explanation of KAN's formulation, readers are referred to the original works [20], [22].

KAN is based on the Kolmogorov-Arnold Representation Theorem [18], which states that which states that any multivariate continuous function can be represented as a finite composition of continuous single-variable functions. In our application, KAN models map generation by simplifying complex transformations into continuous operations, represented by learnable B-spline curves. A key advantage of KAN over traditional MLPs is the use of learnable activation functions, which adaptively enhance feature expressiveness.

Following [22], we reformulate KAN for map generation, applying learnable B-splines to the input tensors without additional entropy regularization during L1 normalization. This two-layer KAN, replacing the MLP, acts as a projection head, which we denote it as K that enhances feature representation before contrastive learning is applied.

a) Query-Selection Attention Mechanism: To address the random query selection in contrastive learning, we incorporate a query-selection attention mechanism, following the approach in [14]. This mechanism computes an attention matrix by evaluating the similarity between a query and its neighboring keys within a fixed window of size $s \times s$. First, we calculate the attention matrix M_{att} by multiplying the reshaped query matrix $Q \in \mathbb{R}^{HW \times C}$ with the key matrix $K \in \mathbb{R}^{HW \times s^2 \times C}$, followed by applying the softmax function, resulting in attention matrix $M_{\text{att}} \in \mathbb{R}^{HW \times s^2}$.

Next, we compute the entropy E of each query as:

$$E(a) = -\sum_{b=1}^{s^2} M_{\text{att}}(a,b) \log M_{\text{att}}(a,b)$$

Here, a and b represent the indices of the query and key,

respectively. We select the top N rows of M_{att} by sorting E in ascending order to form the query-selected attention matrix $M_{\text{q-att}}$. Using these selected N indices, we route values through the value matrix $V \in \mathbb{R}^{HW \times s^2 \times C}$, obtaining a final matrix of relevant queries $R \in \mathbb{R}^{N \times s^2 \times C}$. This process ensures the attention mechanism focuses on spatially relevant features, improving map generation quality.

For training, we use the PatchNCE loss from [26], which is applied at the patch level in a contrastive learning framework. This loss function can be written as:

$$\mathcal{L}_{\text{PatchNCE}}(G, K, X) = \mathbb{E}_{x \sim X}\left[\sum_{l=1}^{L}\sum_{s=1}^{S_l} \ell\left(\hat{z}_l^s, z_l^s, z_l^{S \setminus s}\right)\right] \quad (2)$$

where X denotes the domain of satellite images, $\hat{z}_l^s$ is the query point, z_l^s is the positive point, and $z_l^{S \setminus s}$ are the negative points. L is the total number of layers from which the features are extracted and passed to the projection head K, and S_l is the total number of spatial locations in the l-th layer.

D. Total Loss Functions

To obtain map generation from satellite images with correct depiction of object layers in maps in corresponding location, we used above mention PatchNCE loss, L_{PatchNCE}(G, K, X). Along this, to obtain high quality generation based on human perception, we used Adversarial Loss as presented in [23] constructed under Least Squares Generative Adversarial Networks (LSGANs) principle, initially proposed in [8]. Substantially, we used another PatchNCE loss L_{PatchNCE}(G, K, Y), similar to the previous PatchNCE loss, except that the input domain is Y, instead of X. Hence, the final loss function is the combination of these three loss functions, which we represent as:

$$\begin{aligned}\mathcal{L}_{\text{final}} =& \mathcal{L}_{\text{LSGAN}}(G, D, X, Y) + \quad \lambda_{\text{X}}\mathcal{L}_{\text{PatchNCE}}(G, K, X) + \\ & \lambda_{\text{Y}}\mathcal{L}_{\text{PatchNCE}}(G, K, Y)\end{aligned} \quad (3)$$

IV. Experiments and Results

A. Datasets

a) Aerial Photograph ↔ Maps [15]: This dataset comprises 2,194 unpaired aerial images and map tiles sourced from Google Maps for the training set, all covering regions in and around New York City. The test set includes 2,196 aerial images and corresponding styled map tiles.

b) Styled Maps ↔ Simple Maps [16]: This dataset consists of styled maps and their corresponding simple maps obtained from OpenStreetMap (OSM) vector data at zoom level 15. We used the styled maps as input images and simple, human-readable maps as target images. For the unsupervised task, we utilized 1,740 images for training and 436 for testing. This dataset is included to demonstrate the adaptability and generalizability of our proposed model in generating maps from various input types.

B. Evaluation Metrics

a) Fréchet Inception Distance (FID): We selected the Fréchet Inception Distance (FID) [12] score to evaluate the generative performance of all models. FID is widely used to assess the performance of GANs by calculating the divergence between the distributions of real and generated images in a feature space. This is done by extracting feature vectors using a pre-trained Inception v3 network [29]. FID has proven effective in capturing perceptual differences that align well with human judgment. It is noteworthy that lower FID score equivalents to better generated quality of images.

b) Kernel Inception Distance (KID): The Kernel Inception Distance (KID) is another metric used to evaluate the similarity between real and generated images, based on their feature representations in our setup. Like FID, KID utilizes a pre-trained Inception network to extract feature embeddings from both sets of images. However, instead of assuming a Gaussian distribution as FID does, KID computes the squared Maximum Mean Discrepancy (MMD) between these embeddings using polynomial kernels. This approach makes KID more robust to smaller sample sizes and less prone to bias in comparison to FID. A lower KID score indicates that the generated images are closer in quality to the real images, with better alignment of distributions in the feature space.

c) Peak Signal-to-Noise Ratio (PSNR): PSNR measures image quality by comparing the highest possible pixel value to the noise affecting image fidelity. It is defined as:

$$PSNR = 10 \cdot \log_{10}\left(Max_s^2/MSE\right)$$

where Max_s is the maximum pixel value and MSE is the mean squared error between the original and generated images. Higher PSNR indicates better image quality.

d) Structural Similarity Index (SSIM): SSIM evaluates image quality by comparing luminance, contrast, and structure between original and generated images. It is defined as:

$$SSIM(s, s') = \frac{(2\mu_s\mu_{s'} + C)(2\sigma_{ss'} + C)}{(\mu_s^2 + \mu_{s'}^2 + C)(\sigma_s^2 + \sigma_{s'}^2 + C)}$$

where μ_s and $\mu_{s'}$ are the average luminance values, σ_s^2 and $\sigma_{s'}^2$ are their variances, and $\sigma_{ss'}$ is the covariance. C is a constant to avoid division by zero. Higher SSIM values indicate better image quality.

C. Experimental Environment and Baselines

Our experiments were conducted using Python 3.6.8 and the PyTorch framework for both training and testing. All computations were carried out on a system facilitated with NVIDIA A100-PCI GPUs, each offering 80 GB of HBM2 memory. We utilized CUDA version 12.3 and NVIDIA driver version 545.23.08. The training was carried out over 400 epochs with an initial learning rate of 0.0001, which decayed after the first 200 epochs, using the Adam optimizer. The coefficients were set to $\lambda_{\text{X}} = 1$ and $\lambda_{\text{Y}} = 1$.

1) Baselines: To evaluate the effectiveness of our proposed model, KAN-Attn, we compared it against several advanced GAN-based models for map generation, including CycleGAN [34], AttentionGAN [30], MapCUT [21], and the state-of-the-art HPix [31], specifically on the Aerial Photograph ↔ Maps dataset. These models have demonstrated high performance in this well-utilized dataset, making them suitable baselines for evaluating the capabilities of KAN-Attn.

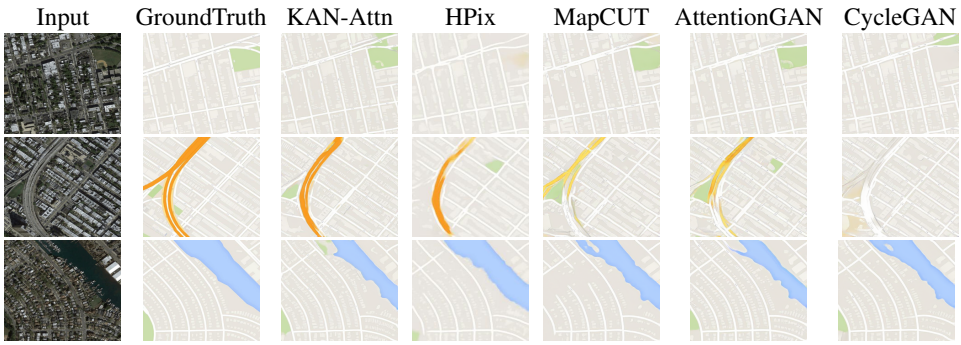

Fig. 3: **Comparative Results of Image Generation from Aerial Photograph ↔ Maps**. Images are reproduced under the Creative Commons Attribution-NonCommercial 4.0 International License for non-commercial research and educational use.

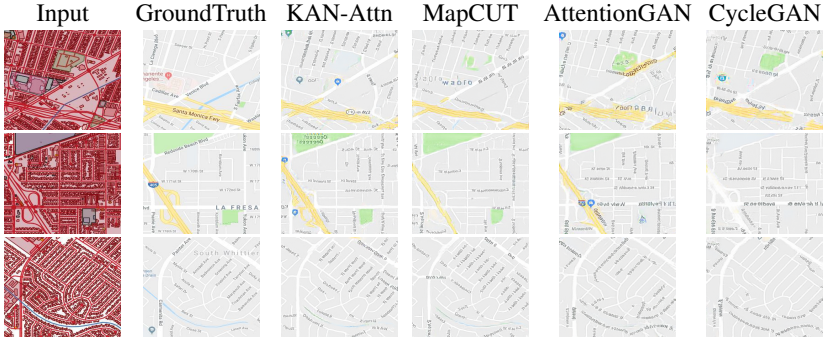

Fig. 4: **Comparative Results of Map Generation from Styled ↔ Simple Maps**. Images reproduced under the MIT License by GeoDS Lab, UW-Madison.

D. Results

The proposed model, KAN-Attn, achieved state-of-the-art performance in generating maps from satellite images, with an FID score of 52.99 and KID of 0.03 on the Aerial Photograph ↔ Maps dataset, as presented in Table 3. This surpasses all existing models, including recent advancement like HPix [31] and MapCUT [21]. In addition to the quantitative evaluation, qualitative results visually demonstrate the superiority of KAN-Attn in generative quality, as depicted in Fig. 3. While KAN-Attn did not outperform state-of-the-art models in SSIM and PSNR, we argue that it maintains state-of-the-art performance in map generation based on perceptual quality, as evidenced by the FID and KID metrics. These metrics, especially FID, are considered the most relevant and reliable for evaluating the generative quality of GANs, as supported by several seminal works [1], [17], [6], [26]. Furthermore, qualitative visual comparisons in Fig. 3 validate the model's superior map image generation quality.

Similarly, KAN-Attn demonstrated similar performance one the Styled Maps ↔ Simple Maps dataset, achieving FID and KID scores of 72.20 and 0.045, respectively, as summarized in Table II. These results also outperform all baseline models. Qualitative results further highlight the model's ability to generate high-quality map images, as shown in Fig. 4.

TABLE I: Quantitative Observation of Results Obtained from All Models in Aerial Photograph ↔ Maps

Model	SSIM	PSNR	FID	KID
CycleGAN	0.63	24.05	67.51	0.09
AttentionGAN	0.71	25.73	72.99	0.13
MapCUT	0.64	24.32	59.37	0.18
HPix	**0.75**	**26.98**	96.71	0.19
KAN-Attn	0.73	25.79	**52.99**	**0.03**

TABLE II: Quantitative Observation of Results Obtained from All Models in Styled Maps ↔ Simple Maps

Model	SSIM	PSNR	FID	KID
CycleGAN	0.48	20.09	74.89	0.053
AttentionGAN	0.47	19.96	74.29	0.052
MapCUT	**0.58**	**20.34**	73.64	0.048
KAN-Attn	0.56	20.01	**72.20**	**0.045**

V. Conclusions and Future Work

Map generation from satellite imagery presents immense potential for advancements in autonomous navigation and post-disaster management, though it faces significant challenges, both in the task itself and in the limitations of generative models. To address these, we proposed KAN-Attn, a novel model leveraging Kolmogorov-Arnold Networks (KAN) as a projection head for generating enriched features, combined with an attention-based mechanism for optimal query selection to enhance contrastive learning. Our model has demonstrated superior performance over advanced baseline models in both quantitative and qualitative evaluations. Importantly, the goal of map generation is to produce human-readable maps, making perceptual quality paramount. We claim state-of-the-art performance based on FID and KID scores, which align closely with human perception, supported by visual results that show better

generative quality.

The innovative use of KAN as a projection head highlights its potential for further contributions to map generation. In future work, we plan to fully integrate KAN into the generator's encoder layers. Along this, given the scarcity of diverse satellite and map datasets, we aim to collect new datasets and evaluate the model's broader applicability.

ACKNOWLEDGMENT

This material is based in part upon work supported by the National Science Foundation under Grant Nos. CNS-2018611 and CNS-1920182 and FDEP grant C-2104. We sincerely thank the reviewers for their valuable comments and suggestions, which helped improve the quality of this paper.

REFERENCES

[1] Andrew Brock. Large scale gan training for high fidelity natural image synthesis. *arXiv preprint arXiv:1809.11096*, 2018.

[2] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In *International conference on machine learning*, pages 1597–1607. PMLR, 2020.

[3] Xu Chen, Songqiang Chen, Tian Xu, Bangguo Yin, Jian Peng, Xiaoming Mei, and Haifeng Li. Smapgan: Generative adversarial network-based semisupervised styled map tile generation method. *IEEE Transactions on Geoscience and Remote Sensing*, 59(5):4388–4406, 2020.

[4] Yuanqi Chen, Xiaoming Yu, Shan Liu, Wei Gao, and Ge Li. Zero-shot unsupervised image-to-image translation via exploiting semantic attributes. *Image and Vision Computing*, 124:104489, 2022.

[5] Minjong Cheon. Kolmogorov-arnold network for satellite image classification in remote sensing. *arXiv preprint arXiv:2406.00600*, 2024.

[6] Yunjey Choi, Youngjung Uh, Jaejun Yoo, and Jung-Woo Ha. Stargan v2: Diverse image synthesis for multiple domains. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 8188–8197, 2020.

[7] Swetava Ganguli, Pedro Garzon, and Noa Glaser. Geogan: A conditional gan with reconstruction and style loss to generate standard layer of maps from satellite images. *arXiv preprint arXiv:1902.05611*, 2019.

[8] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *Advances in neural information processing systems*, 27, 2014.

[9] Junlin Han, Mehrdad Shoeiby, Lars Petersson, and Mohammad Ali Armin. Dual contrastive learning for unsupervised image-to-image translation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 746–755, 2021.

[10] Yashas Hariprasad, SS Iyengar, and N Subramanian. Deepfake video detection using lip region analysis with advanced artificial intelligence based anomaly detection technique.

[11] Yashas Hariprasad, KJ Latesh Kumar, L Suraj, and SS Iyengar. Boundary-based fake face anomaly detection in videos using recurrent neural networks. In *Proceedings of SAI Intelligent Systems Conference*, pages 155–169. Springer, 2022.

[12] Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter. Gans trained by a two time-scale update rule converge to a local nash equilibrium. *Advances in neural information processing systems*, 30, 2017.

[13] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. Multilayer feedforward networks are universal approximators. *Neural networks*, 2(5):359–366, 1989.

[14] Xueqi Hu, Xinyue Zhou, Qiusheng Huang, Zhengyi Shi, Li Sun, and Qingli Li. Qs-attn: Query-selected attention for contrastive learning in i2i translation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 18291–18300, 2022.

[15] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

[16] Yuhao Kang, Song Gao, and Robert E Roth. Transferring multiscale map styles using generative adversarial networks. *International Journal of Cartography*, 5(2-3):115–141, 2019.

[17] Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 4401–4410, 2019.

[18] A.N. Kolmogorov. *On the representation of continuous functions of several variables by superpositions of continuous functions of a smaller number of variables*. American Mathematical Society, 1961.

[19] Yuanyuan Liu, Wenbin Wang, Fang Fang, Lin Zhou, Chenxing Sun, Ying Zheng, and Zhanlong Chen. Cscgan: Conditional scale-consistent generation network for multi-level remote sensing image to map translation. *Remote Sensing*, 13(10):1936, 2021.

[20] Ziming Liu, Yixuan Wang, Sachin Vaidya, Fabian Ruehle, James Halverson, Marin Soljačić, Thomas Y Hou, and Max Tegmark. Kan: Kolmogorov-arnold networks. *arXiv preprint arXiv:2404.19756*, 2024.

[21] Arpan Mahara and Naphtali D Rishe. Generative adversarial model equipped with contrastive learning in map synthesis. In *Proceedings of the 2024 6th International Conference on Image Processing and Machine Vision*, pages 107–114, 2024.

[22] Arpan Mahara, Naphtali D Rishe, and Liangdong Deng. The dawn of kan in image-to-image (i2i) translation: Integrating kolmogorov-arnold networks with gans for unpaired i2i translation. *arXiv preprint arXiv:2408.08216*, 2024.

[23] Xudong Mao, Qing Li, Haoran Xie, Raymond YK Lau, Zhen Wang, and Stephen Paul Smolley. Least squares generative adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pages 2794–2802, 2017.

[24] Tushar Nayan, Qiming Guo, Mohammed Al Duniawi, Marcus Botacin, Selcuk Uluagac, and Ruimin Sun. {SoK}: All you need to know about {On-Device}{ML} model extraction-the gap between research and practice. In *33rd USENIX Security Symposium (USENIX Security 24)*, pages 5233–5250, 2024.

[25] Aaron van den Oord, Yazhe Li, and Oriol Vinyals. Representation learning with contrastive predictive coding. *arXiv preprint arXiv:1807.03748*, 2018.

[26] Taesung Park, Alexei A Efros, Richard Zhang, and Jun-Yan Zhu. Contrastive learning for unpaired image-to-image translation. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part IX 16*, pages 319–345. Springer, 2020.

[27] Jieqiong Song, Hao Chen, Chun Du, and Jun Li. Semi-mapgen: Translation of remote sensing image into map via semisupervised adversarial learning. *IEEE Transactions on Geoscience and Remote Sensing*, 61:1–19, 2023.

[28] Jieqiong Song, Jun Li, Hao Chen, and Jiangjiang Wu. Mapgen-gan: A fast translator for remote sensing image to map via unsupervised adversarial learning. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14:2341–2357, 2021.

[29] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2818–2826, 2016.

[30] Hao Tang, Hong Liu, Dan Xu, Philip HS Torr, and Nicu Sebe. Attentiongan: Unpaired image-to-image translation using attention-guided generative adversarial networks. *IEEE transactions on neural networks and learning systems*, 34(4):1972–1987, 2021.

[31] Aditya Taparia and Keshab Nath. Hpix: Generating vector maps from satellite images. *arXiv preprint arXiv:2407.13680*, 2024.

[32] Wenjia Wang, Seyed Masoud Sadjadi, and Naphtali Rishe. Curse of feature selection: a comparison experiment of ddos detection using classification techniques. In *2022 IEEE Intl Conf on Parallel Distributed Processing with Applications, Big Data Cloud Computing, Sustainable Computing Communications, Social Computing Networking (ISPA/BDCloud/SocialCom/SustainCom)*, pages 262–269, 2022.

[33] Jian Xu, Xiaowen Zhou, Chaolin Han, Bing Dong, and Hongwei Li. Sam-gan: supervised learning-based aerial image-to-map translation via generative adversarial networks. *ISPRS International Journal of Geo-Information*, 12(4):159, 2023.

[34] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pages 2223–2232, 2017.