УДК: 004.932

Применение сетей Колмогорова-Арнольда для решения задач компьютерного зрения

К.А. Бунин[1,2] А.В. Шешкус[2]

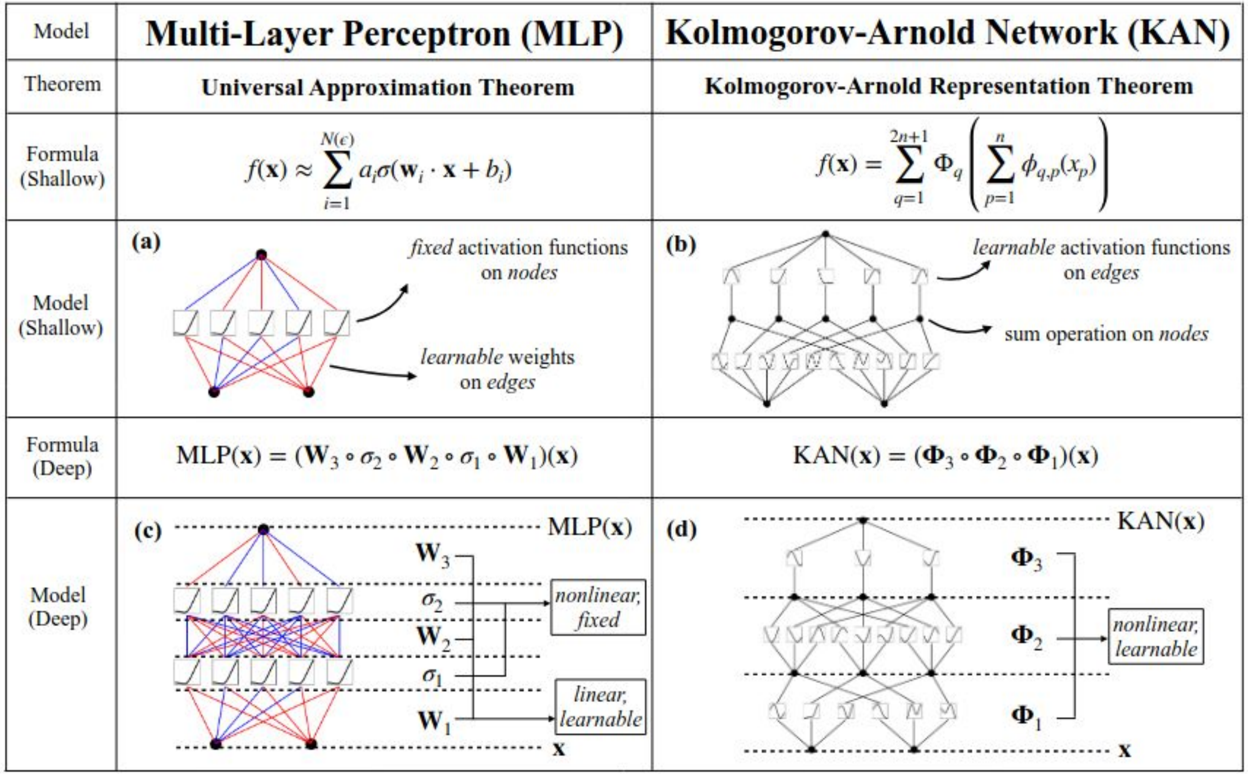
[1]МФТИ [2]Smart Engines Service LLC

План доклада

1) Обзор KAN
2) Литобзор
3) Направление исследования
4) Текущие результаты
5) Дальнейшее исследование

Обзор KAN

Основная идея KAN



KAN vs MLP

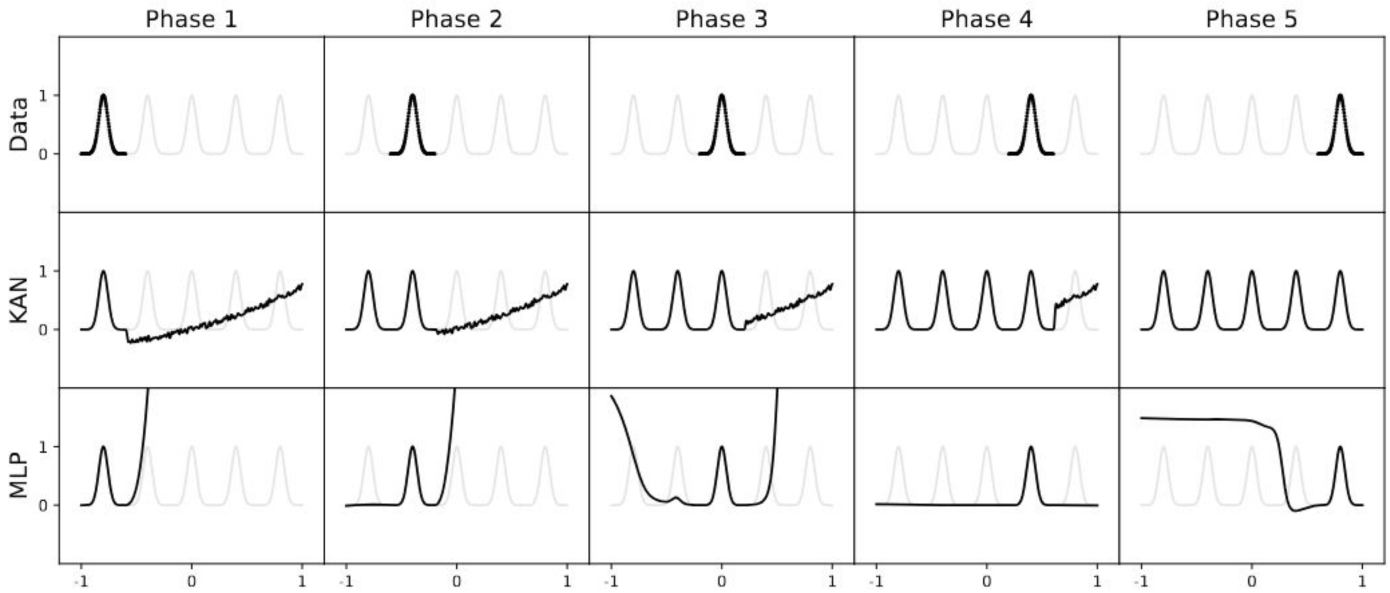


Figure 3.4: A toy continual learning problem. The dataset is a 1D regression task with 5 Gaussian peaks (top row). Data around each peak is presented sequentially (instead of all at once) to KANs and MLPs. KANs (middle row) can perfectly avoid catastrophic forgetting, while MLPs (bottom row) display severe catastrophic forgetting.

KAN в задачах computer vision

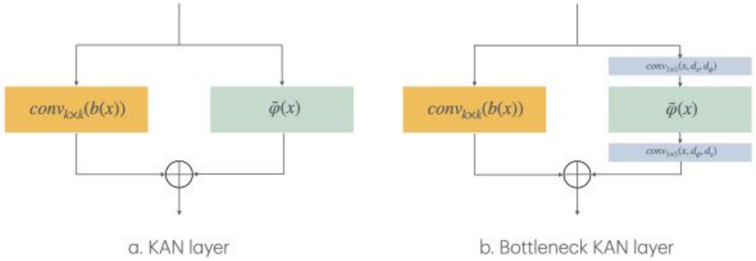


Figure 1: KAN Convolution (left) and Bottleneck KAN Convolution (right). The main difference between these two types of layers is a encoder-decoder convolutional layers on the right data stream.

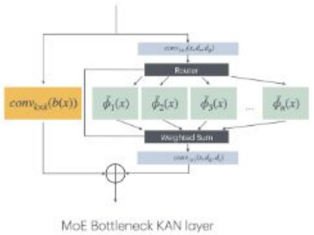


Figure 2: Bottleneck Kolmogorov-Arnold Convolutional Mixture of Experts. The router and experts are placed between bottleneck convolutions, and each expert is a $\tilde{\varphi}$ set of univariate functions. We use sparsely-gated mixture-of-experts [15].

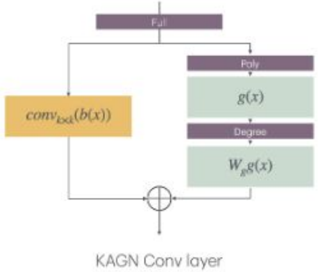


Figure 3: Possible dropout layer placement inside KAGN Convolution layer: Full - before the layer, Poly - before computing Gram basis, and Degree - before weighted sum of previously computed basis.

Литобзор

https://docs.google.com/document/d/1BmdW4PVTmEM5WA25ypgqXOwJL_42EFL2BYlEtCww5f4/edit?tab=t.0

Направление исследования

KOLMOGOROV-ARNOLD CONVOLUTIONS: DESIGN PRINCIPLES AND EMPIRICAL STUDIES

A PREPRINT

Ivan Drokin
Deep Learning Researcher
Seath the Scaleless Research Group
ivan.s.drokin@gmail.com

July 2, 2024

KAN not Work: Investigating the Applicability of Kolmogorov-Arnold Networks in Computer Vision

Yueyang Cang[1*] Yuhang Liu[1*] Shi Li[1†]
[1]Tsinghua University, Beijing, China
cangyy23@mails.tsinghua.edu.cn, yh-liu23@mails.tsinghua.edu.cn, shilits@tsinghua.edu.cn

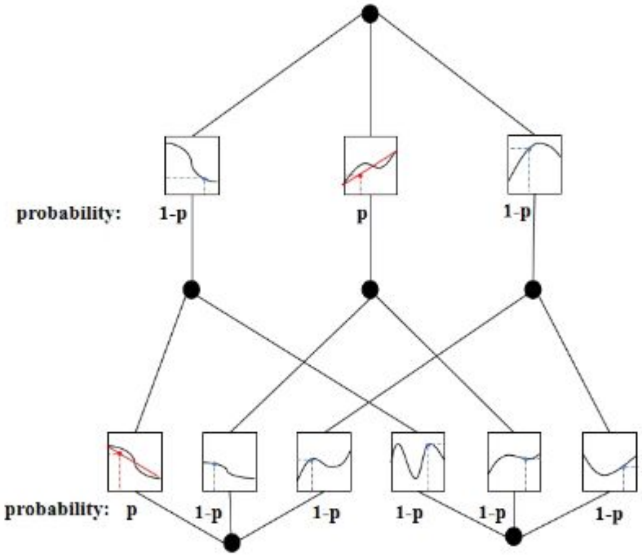


Table 2. **Experimental Model Configurations**. The baseline model (CNN+MLP) serves as a control for comparison with the CKAN and KAN-modified models.

Model	20%	40%	60%	80%	100%
CNN+MLP	62.31%	67.26%	69.93%	71.33%	71.96%
CKAN+CNN+MLP	61.82%	67.05%	69.03%	70.91%	71.88%
CNN+KAN	64.02%	68.15%	70.41%	72.13%	72.58%

Table 3. **Performance of Different Models on CIFAR-10 with Varying Dataset Sizes**. Each column shows the accuracy (%) of each model as the dataset size increases from 20% to 100%.

Model	10%	20%	30%	40%	50%
CNN+MLP	69.13%	67.93%	65.38%	63.88%	62.87%
CKAN+CNN+MLP	69.59%	66.73%	64.19%	62.56%	58.43%
CNN+KAN	69.95%	67.81%	64.33%	62.84%	59.20%

Table 4. **Performance of Different Models with Increasing Label Noise**. Each column represents the accuracy (%) as the label noise increases from 10% to 50%.

U-Net-bin: hacking the document image binarization contest

P.V. Bezmaternykh[1,2], D.A. Ilin [1], D.P. Nikolaev [1, 3]
[1] Smart Engines Service LLC, 117312, Moscow, Russia,
[2] Federal Research Center "Computer Science and Control" of RAS, 117312, Moscow, Russia,
[3] Institute for Information Transmission Problems of RAS, 127051, Moscow, Russia

a) b)

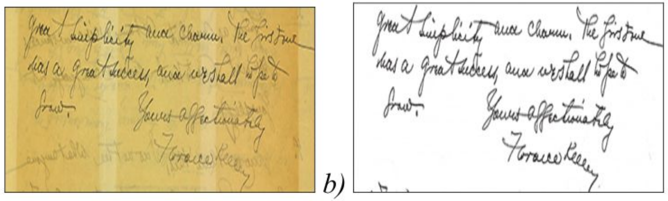


Fig. 1. Sample of document binarization application: (a) source image, (b) binarization result

Brief	Score	FM	Fps	PSNR	DRD
1) U-Net (Proposed method)	**309**	**91.01**	**92.86**	**18.28**	**3.40**
2) FCN (VGGNet)	455	89.67	91.03	17.58	4.35
3) Ensemble (3 DSN)	481	89.42	91.52	17.61	3.56
4) Ensemble (5 FCN, no postprocessing)	529	86.05	90.25	17.53	4.52
5) Ensemble (FCN, with postprocessing)	566	83.76	90.35	17.07	4.33
6) FCN	608	88.37	89.59	17.10	4.94
7) Howe based	635	89.17	89.88	17.85	5.66
Otsu	-	77.73	77.89	13.85	15.5
Sauvola	-	77.11	84.10	14.25	8.85

U-KAN Makes Strong Backbone for Medical Image Segmentation and Generation

Chenxin Li[1*], Xinyu Liu[1*], Wuyang Li[1*], Cheng Wang[1*],
Hengyu Liu[1], Yifan Liu[1], Zhen Chen[2], Yixuan Yuan[1],

[1]The Chinese University of Hong Kong, [2]CAIR, HKISI-CAS

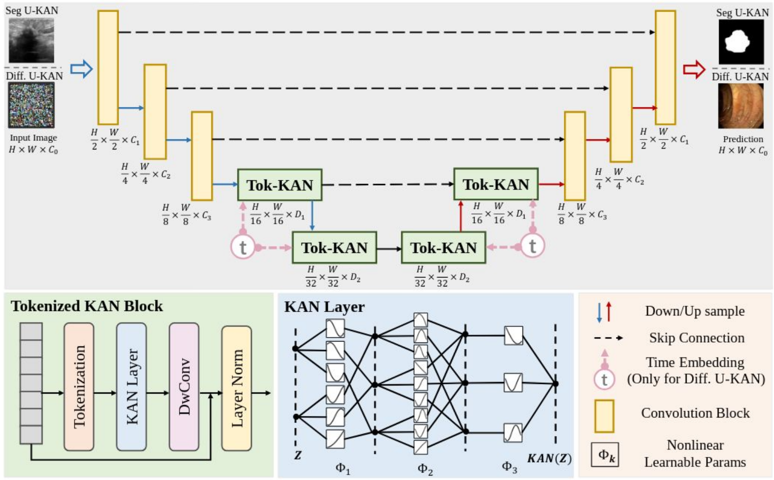
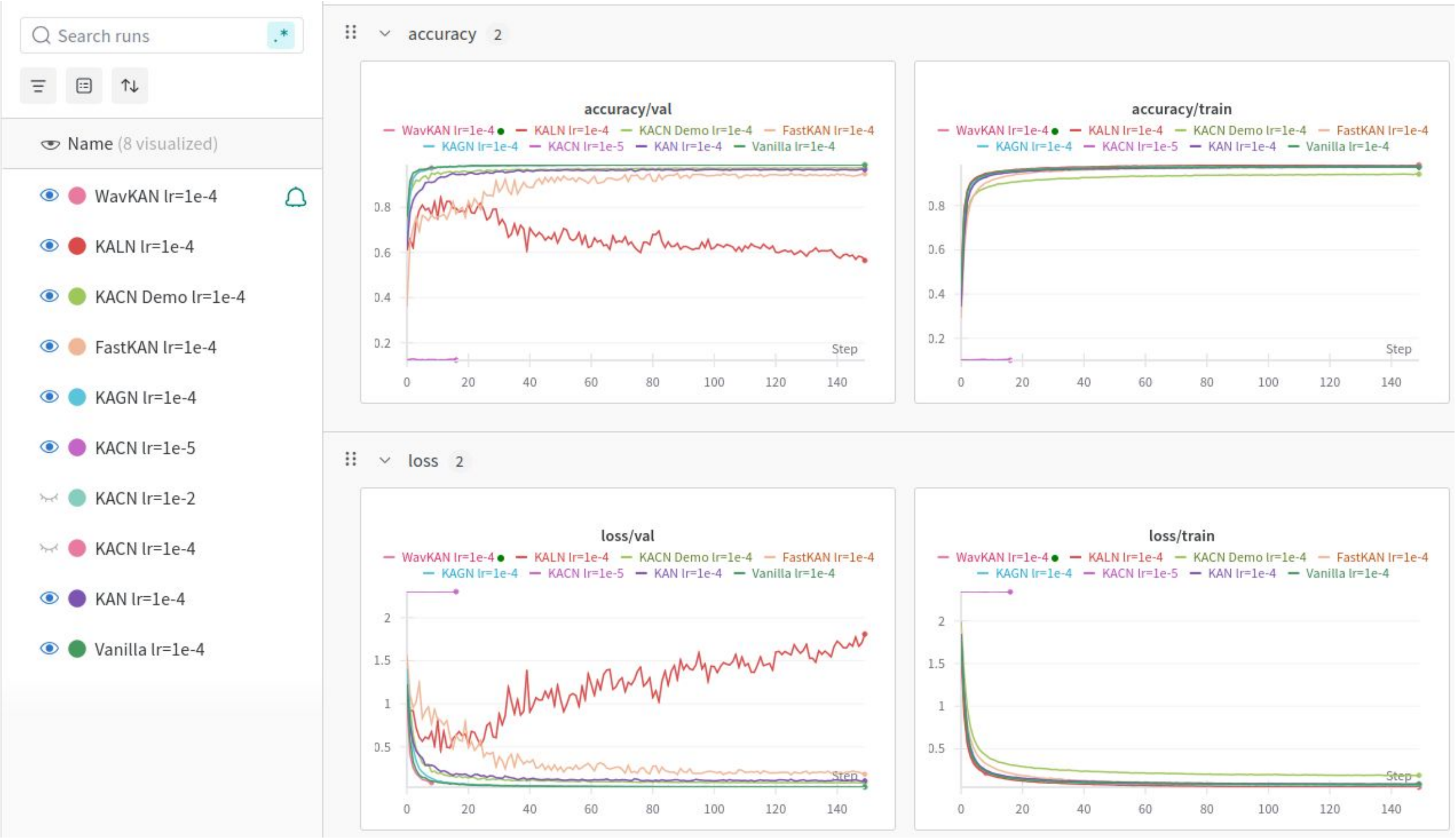


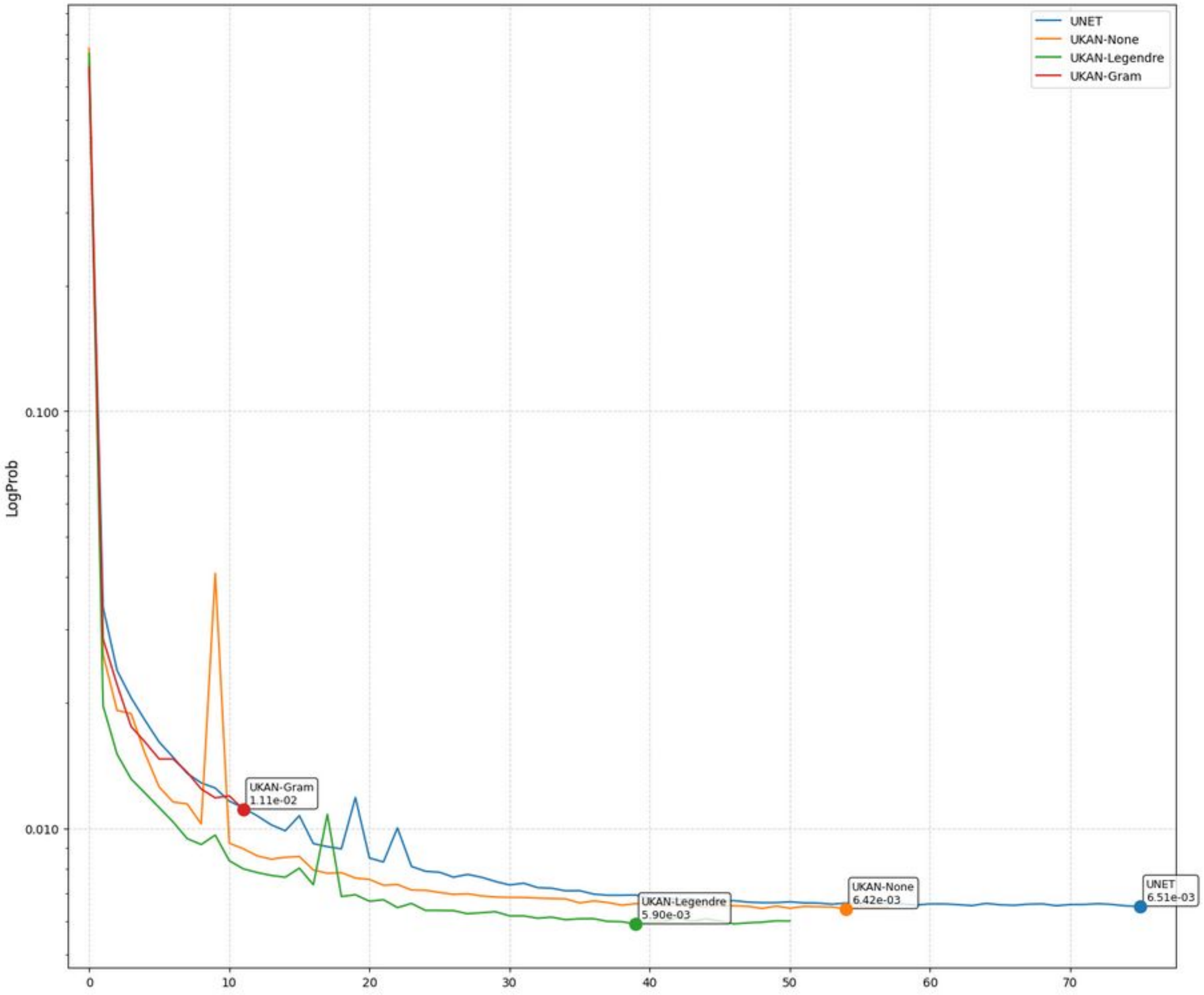
Figure 1: Overview of U-KAN pipeline. After feature extraction by several convolution blocks in Convolution Phrase, the intermediate maps are tokenized and processed by stacked Tok-KAN blocks in Tokenized KAN Phrase. The time embedding is only injected into the KAN blocks when applied for Diffusion U-KAN.

Table 5: Ablation studies on using KAN layers against MLPs. The default setup is denoted.

KAN vs. MLP	IoU↑	F1↑	Gflops
KAN×3	65.26	78.75	14.02
MLP+KAN+KAN	64.12	77.86	14.29
KAN+MLP+KAN	63.82	77.58	14.29
KAN+KAN+MLP	64.30	77.95	14.29
MLP×3	63.49	77.07	14.84

Текущие результаты


Search runs
Name (8 visualized)
WavKAN lr=1e-4
KALN lr=1e-4
KACN Demo lr=1e-4
FastKAN lr=1e-4
KAGN lr=1e-4
KACN lr=1e-5
KACN lr=1e-2
KACN lr=1e-4
KAN lr=1e-4
Vanilla lr=1e-4
accuracy 2
accuracy/val
accuracy/train
loss 2
loss/val
loss/train
Step



UNET
UKAN-None
UKAN-Legendre
UKAN-Gram
LogProb
0.100
0.010
0
10
20
30
40
50
60
70
UKAN-Gram
1.11e-02
UKAN-Legendre
5.90e-03
UKAN-None
6.42e-03
UNET
6.51e-03

Результаты

1) Имплементирован KAN с различными базисами
2) Сделаны эксперименты на датасетах MNIST, CIFAR, DIBCO
3) Улучшена UNET-архитектура для DIBCO

Дальнейшее исследование

1) Исследование Conv KAN
2) Исследование Wavelet KAN
3) Квантование KAN