

# **Использование слоёв KAN-свёртки в задаче сегментации изображений и классификации изображений.**

Автор: Игорь Петрович

**Аннотация:** За последнее время нейронные сети архитектуры KAN[1] достигли значительных результатов[2,3] в компьютерном зрении. Смысл данной работы использованием различных архитектур(в нашем случае - FCN) моделей компьютерного зрения со встроенными KAN слоями[2]. Эта работа доказывает - то что модели построенные на KAN архитектуре, могут не уступать классическим архитектурам в некоторых задачах компьютерного зрения.

**Вступление:** Цель данной работы взять архитектуру модели компьютерного зрения, и встроить в неё слои KAN, после провести эксперименты на различных датасетах. За архитектуру была взята Fully Convolutional Network. За слои KAN использовались conv\_KAN слои из статьи [2]. Как мы впоследствии можем увидеть архитектура построенная на KAN\_conv слоях, неплохо показала себя в задачах классификации, и сегментации изображений.

**Предыдущие работы:**

**Kolmogorov-Arnold Network(KAN):**

Архитектура нейронной сети предложенная в статье [1]. Основана на теореме Колмогорова-Арнольда[4]. Основная идея KAN - состоит в замене соединений (весов(числовые переменные)) между слоями B-сплайнами. В статье также упоминается что KAN более точны, и интерпретируемы чем MLP.

**Fully-Convolutional-Network:**

Архитектура нейронной сети созданная для сегментации изображений. В архитектуре используется только слои свёртки, пулинга, и повышения разрешения. Сеть состоит из пути понижения частоты дискретизации, используемого для извлечения, и интерпретации контекста, и пути повышения частоты дискретизации, который позволяет локализовать данные. Впервые упоминается в статье [5].

**KAN\_convolutional layer:**

Впервые введён в статье [2]. Представляет собой слой свёртки в фильтре которого вместо весов используется В-сплайны.

### Метод и архитектура:

В архитектуре модели было 6 слоёв. Сначала, шли 2 пары conv-maxpool слоёв. Позже карты признаков передавались в 2 conv слоя. Размеры фильтра у всех conv слоёв(кроме последнего) 3x3, у последнего слоя 1x1.размеры свертки maxpool 2x2.Модель обучалась на датасете Pascal Voc(для задачи сегментации),и на датасетах CIFAR-10, CIFAR-100(для задачи классификации).

### Эксперименты и итоги:

датасет	CIFAR-10	CIFAR-100	Pascal Voc
Train_loss	1.361	4.608	0.1854
Test_accuracy/loss(на 10000 изображений/на 1000)	Accuracy:54%	Accuracy:1%	Loss:2.0998

В отличие от MLP, KAN требуется намного меньше слоёв, и нейронов ввиду формулировки теоремы, и эмпирического опыта, который мы получили при экспериментах. Также как и MLP, KAN подвержены переобучению поэтому архитектуру модели нужно изменять под сложность задачи и кол-во данных. (Примеры работы модели в задаче сегментации изображений в приложении А).

### Вывод

Исходя из данных полученных при экспериментах мы можем сказать что KAN и в частности KAN\_conv может хорошо себя показать в задачах компьютерного зрения и может конкурировать с классическими архитектурами.При этом имея намного меньше слоёв и параметров чем архитектуры основанные на MLP.Мы уверены что KAN может себя хорошо показать и в других архитектурах и задачах[3,6].Весь код на основе которого проводились эксперименты будет публиковаться на:

[https://github.com/companys1234/FCN\\_KAN](https://github.com/companys1234/FCN_KAN)

## Список литературы:

[1]- <https://arxiv.org/pdf/2404.19756>, KAN: Kolmogorov-Arnold Networks

[2]- <https://arxiv.org/pdf/2406.13155>, Convolutional Kolmogorov–Arnold Networks(слои conv\_KAN заимствованы отсюда - <https://github.com/AntonioTepsich/Convolutional-KANs>)

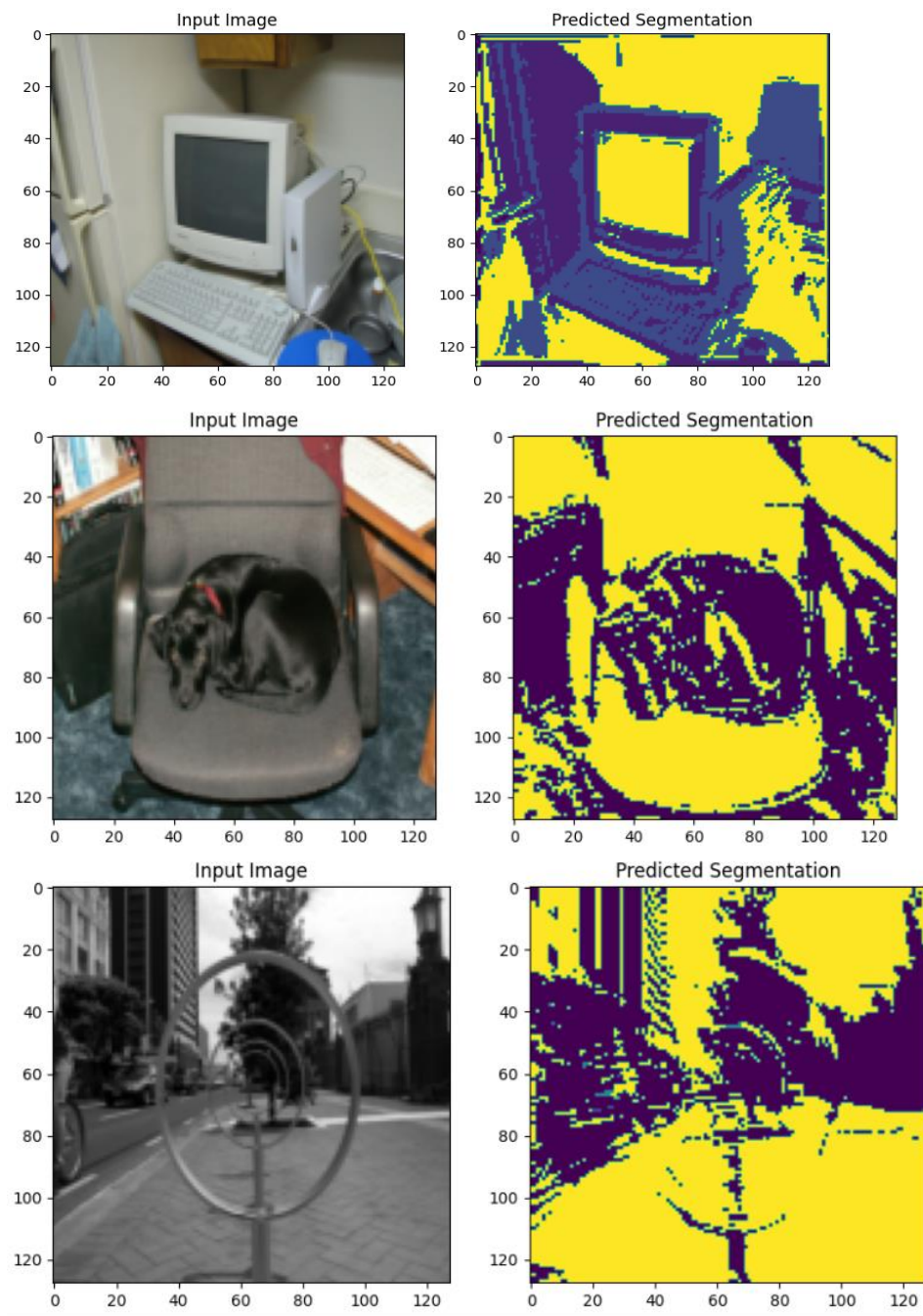
[3]- <https://arxiv.org/pdf/2406.02918>, U-KAN Makes Strong Backbone for Medical Image Segmentation and Generation

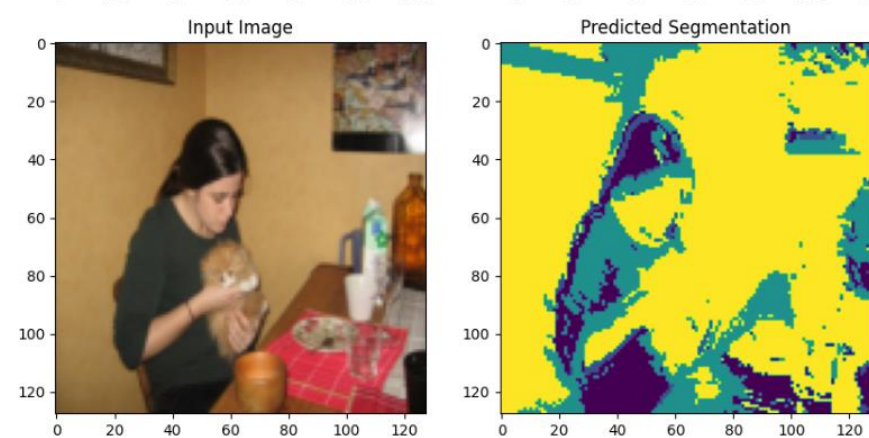
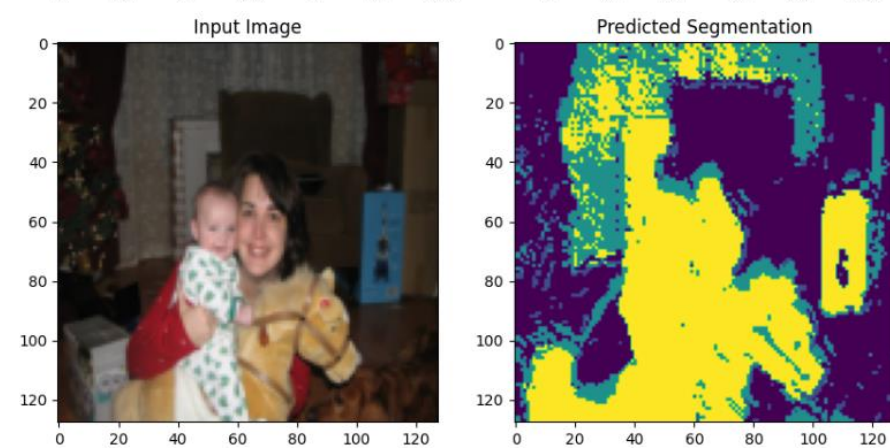
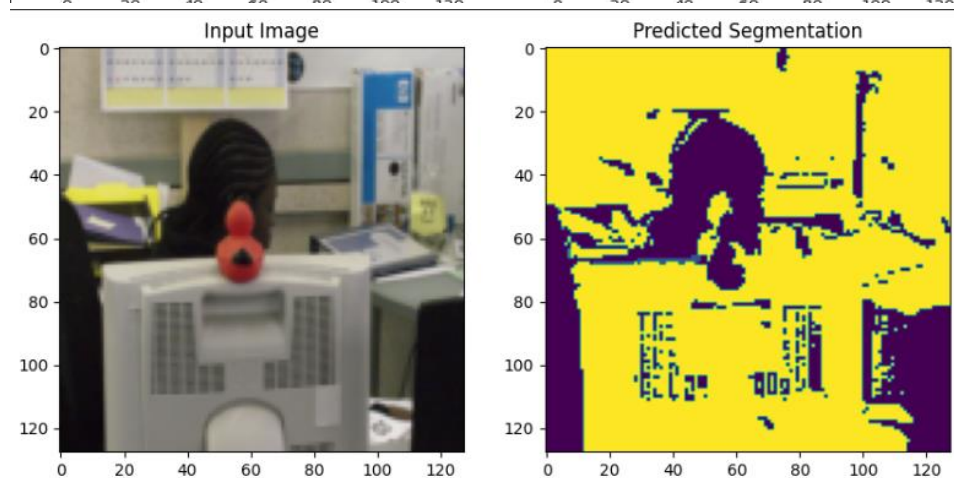
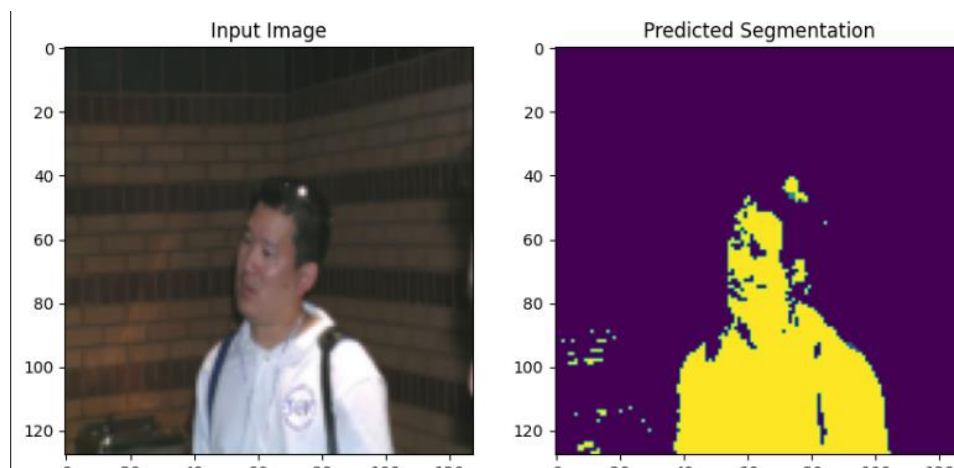
[4][https://ru.wikipedia.org/wiki/%D0%A2%D0%B5%D0%BE%D1%80%D0%B5%D0%BC%D0%B0\\_%D0%9A%D0%BE%D0%BB%D0%BC%D0%BE%D0%B3%D0%BE%D1%80%D0%BE%D0%B2%D0%B0\\_%E2%80%94%D0%90%D1%80%D0%BD%D0%BE%D0%BB%D1%8C%D0%B4%D0%B0](https://ru.wikipedia.org/wiki/%D0%A2%D0%B5%D0%BE%D1%80%D0%B5%D0%BC%D0%B0_%D0%9A%D0%BE%D0%BB%D0%BC%D0%BE%D0%B3%D0%BE%D1%80%D0%BE%D0%B2%D0%B0_%E2%80%94%D0%90%D1%80%D0%BD%D0%BE%D0%BB%D1%8C%D0%B4%D0%B0), Теорема Колмогорова — Арнольда

[5]- <https://arxiv.org/pdf/1411.4038>, Fully Convolutional Networks for Semantic Segmentation

[6]- <https://arxiv.org/pdf/2406.09087>, Suitability of KANs for Computer Vision: A preliminary investigation

## Приложение А:





Рисунки 1-7. Визуализация работы модели на датасете Pascal Voc.