

**ІІТМО**

# **Открытая библиотека методов анализа и генерации векторной графики**

Студенты: Пименов А. В., Баженов Е. А., Жарский И. А.  
Научный руководитель: Ефимова В. А.

Санкт-Петербург, 2024

# Актуальность

	Векторная графика	Растровая графика
Плюсы	<ul style="list-style-type: none"><li>- Масштабируемость</li><li>- Высокое качество изображения</li><li>- Сравнительно небольшой вес</li></ul>	<ul style="list-style-type: none"><li>- Удобство использования</li><li>- Простота создания и редактирования</li><li>- Много решений для обработки</li></ul>
Минусы	<ul style="list-style-type: none"><li>- Не везде поддерживается</li><li>- Сложно редактировать</li><li>- Мало решений для обработки</li></ul>	<ul style="list-style-type: none"><li>- Не масштабируема</li><li>- Квадратичная зависимость веса от размера</li></ul>

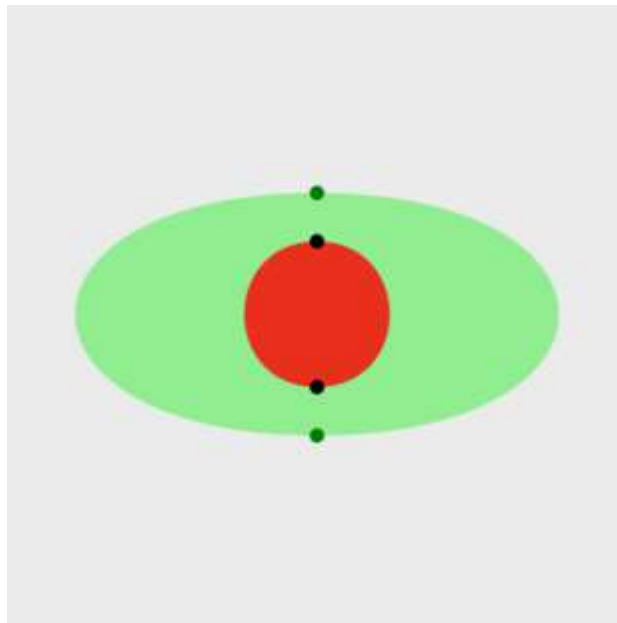


Растровое изображение

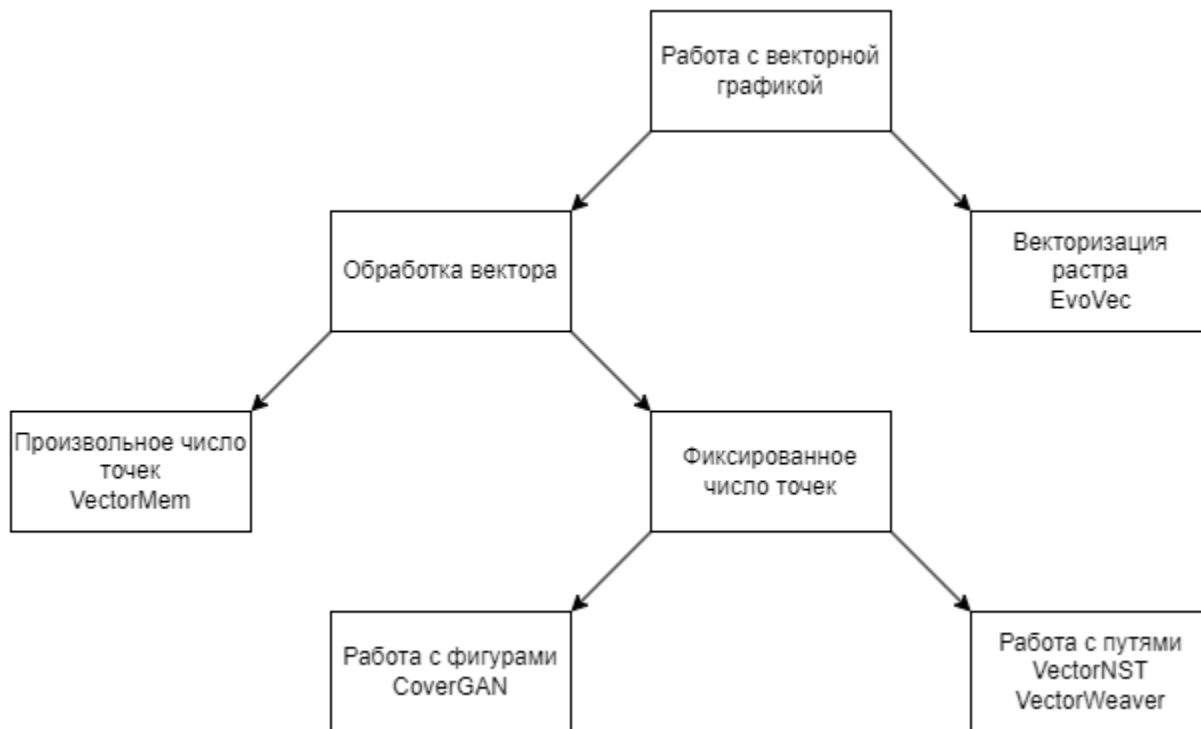
Векторное изображение

# Пример векторного изображения

```
<svg height="256" width="256"
  version="1.1" xmlns="http://www.w3.org/2000/svg">
  <rect fill="rgb(235,235,235)" height="256.0" width="256.0"/>
  <path d="M 128 78
    C 261 78 261 178 128 178
    C -5 178 -5 78 128 78"
    fill="lightgreen" opacity="1.0"/>
  <circle cx="128" cy="78" r="3" fill="green"/>
  <circle cx="128" cy="178" r="3" fill="green"/>
  <path d="M 128 98
    C 168 98 168 158 128 158
    C 88 158 88 98 128 98"
    fill="red" opacity="0.8"/>
  <circle cx="128" cy="98" r="3" fill="black"/>
  <circle cx="128" cy="158" r="3" fill="black"/>
</svg>
```

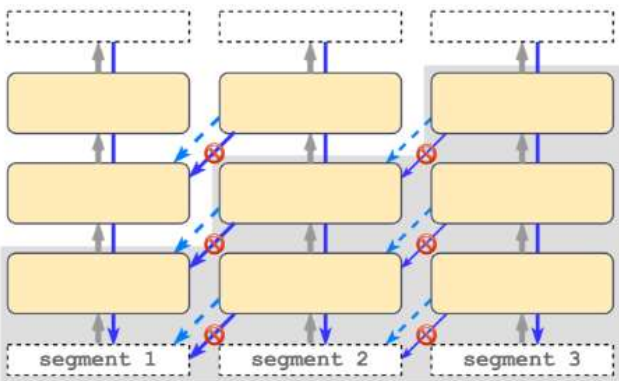


# Варианты обработки векторного изображения

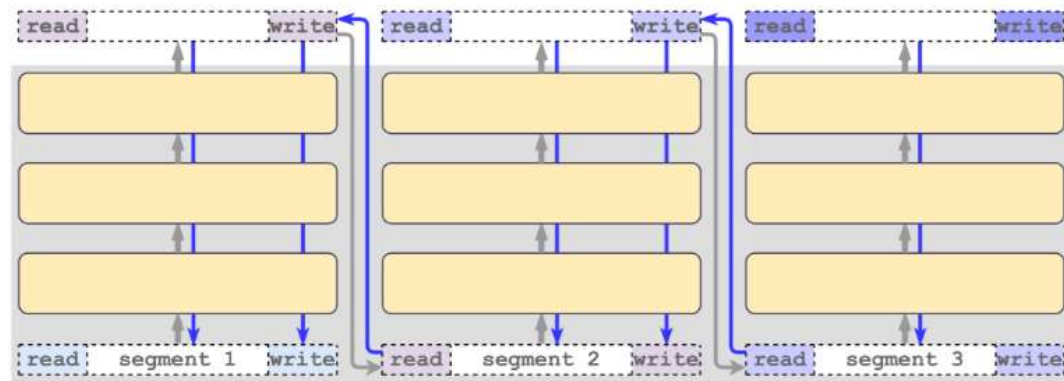


# VectorMem. Трансформеры памяти

### Transformer-XL



### Recurrent Memory Transformer



Transformer layer Attention to cached states Effective context Forward pass Gradients Stop Gradient

Bessonov, Arkadii & Staroverov, Alexey & Zhang, Huzhenyu & Kovalev, Alexey & Yudin, Dmitry & Panov, Aleksandr. (2023). Recurrent Memory Decision Transformer.

# Проблема обычных трансформеров памяти

## Недостатки:

- Недостаточно используемого контекста для обработки сверхвысокоразмерных изображений
- Контекст из начала последовательности данных постепенно “затухает”



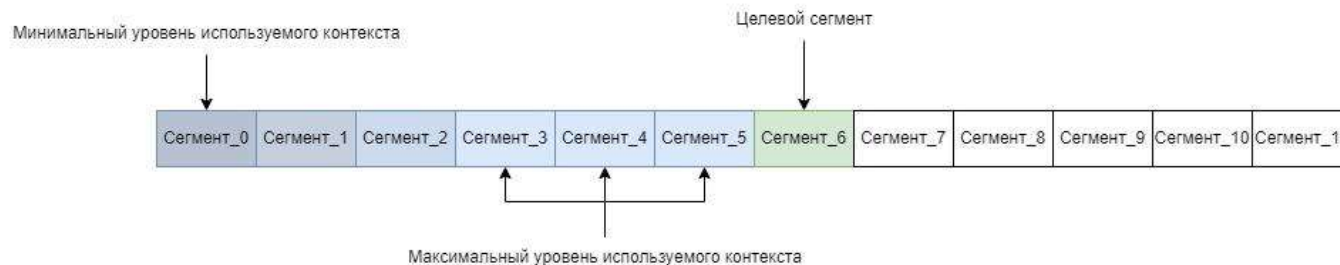
# VectorMem. Список контекста

## Достоинства:

- Расширяет область используемого контекста
- Контекст медленнее затухает
- Полезен на данных с небольшим количеством сегментов

## Недостатки:

- Слабо помогает с проблемой затухания на больших данных



# Сравнение метрик

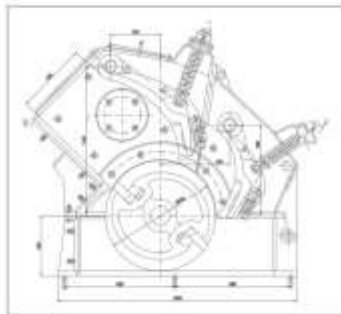
Название метода	Точность обнаружения плагиата	Точность разделения плагиата
Растровый метод	79,4%	5,6%
Сверточные нейронные сети	87%	78%
Рекуррентный трансформер памяти	89%	63%
VectorMem	91%	88%



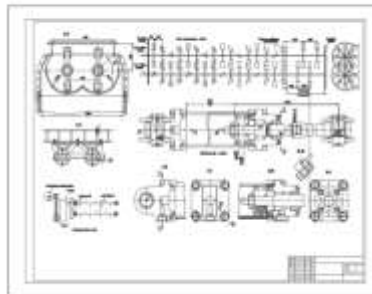
# VectorMem. Примеры поиска плагиата



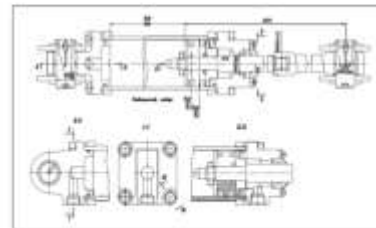
Имя оригинала: 027 - Дробилка центробежная - СМД-05.dwg  
Рабочие чертежи(027027)  
Обнаружен Тух



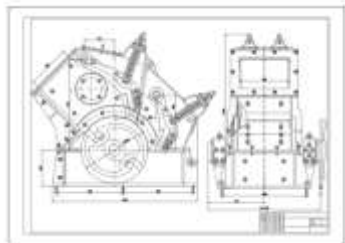
Имя оригинала: 099 - Бетонномешалка С-545.dwg



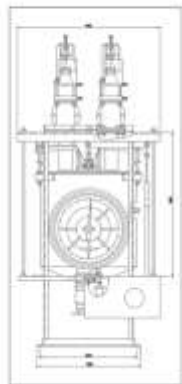
Имя оригинала: 099 - Бетонномешалка С-545.dwg  
Рабочие чертежи(099099)  
Обнаружен Тух



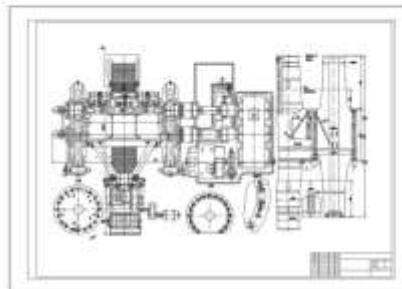
Имя оригинала: 027 - Дробилка центробежная - СМД-05.dwg



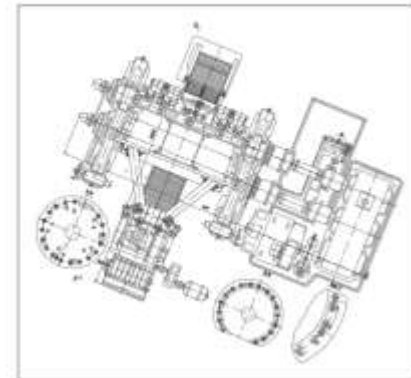
Имя оригинала: 070 - Дозатор.dwg  
Рабочие чертежи(070070)  
Обнаружен Тух



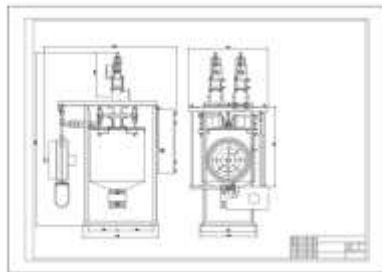
Имя оригинала: 104 - Валцы стальные (Лист 2).dwg



Имя оригинала: 104 - Валцы стальные (Лист 2).dwg  
Рабочие чертежи(104104)  
Обнаружен Тух

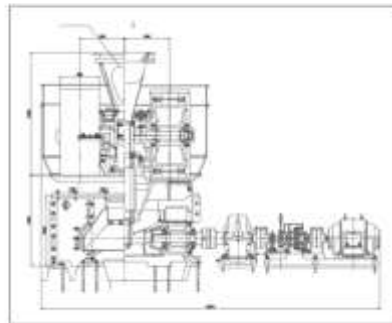


Имя оригинала: 070 - Дозатор.dwg

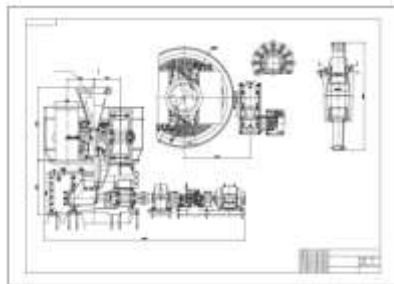


# VectorMet. Примеры поиска плагиата

Имя файла: Бетонный насос CM-365\_dlg.dwg  
Расстояние: 1000(0.0000)  
Обнаружен: True



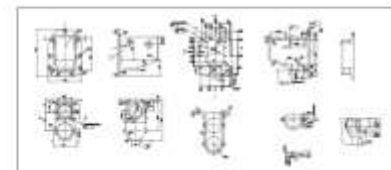
Имя файла: Бетонный насос CM-365.dwg



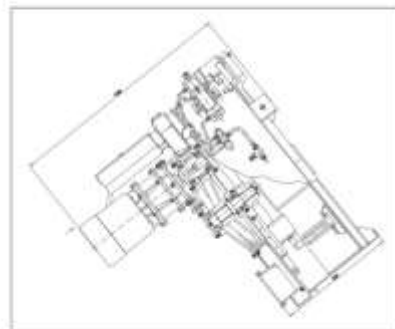
Имя файла: Дозатор.dwg



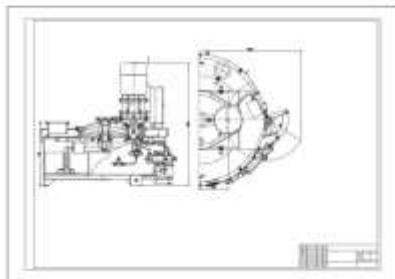
Имя файла: Дозатор\_dlg.dwg  
Расстояние: 1000(0.0000)  
Обнаружен: True



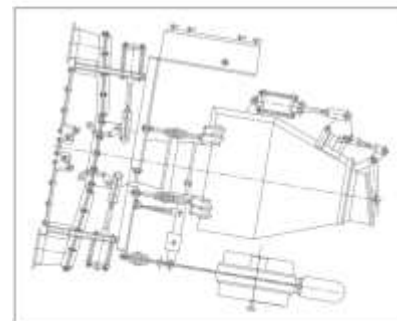
Имя файла: Бетонно-цементный принудительного действия С-773\_dlg.dwg  
Расстояние: 1000(0.0416)  
Обнаружен: True



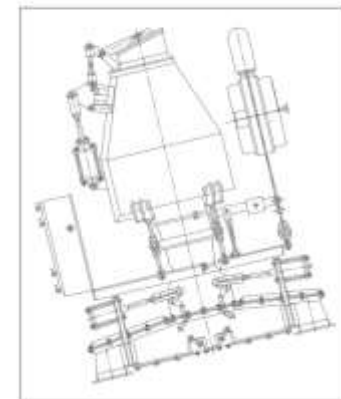
Имя файла: Бетонно-цементный принудительного действия С-773.dwg



Имя файла: Дозатор.dwg



Имя файла: Дозатор\_dlg.dwg  
Расстояние: 1000(0.0000)  
Обнаружен: True



## Преимущества использования:

1. Может обрабатывать векторные изображения любого размера
2. Возможность выбора нужного варианта метода и адаптации параметров для конкретной задачи
3. Снижение потребления ресурсов на хранение и обработку изображений
4. Возможность обработки векторных изображений в нативном виде
5. Возможность использования более гибкого анализа за счет обработки только нужных объектов

## Недостатки использования:

1. Модель необходимо обучать с нуля
2. Требуется больше ресурсов для обучения

# CoverGAN. О сервисе

## Входные параметры:

1. Трек музыкальной композиции.
2. Вызываемая эмоция.

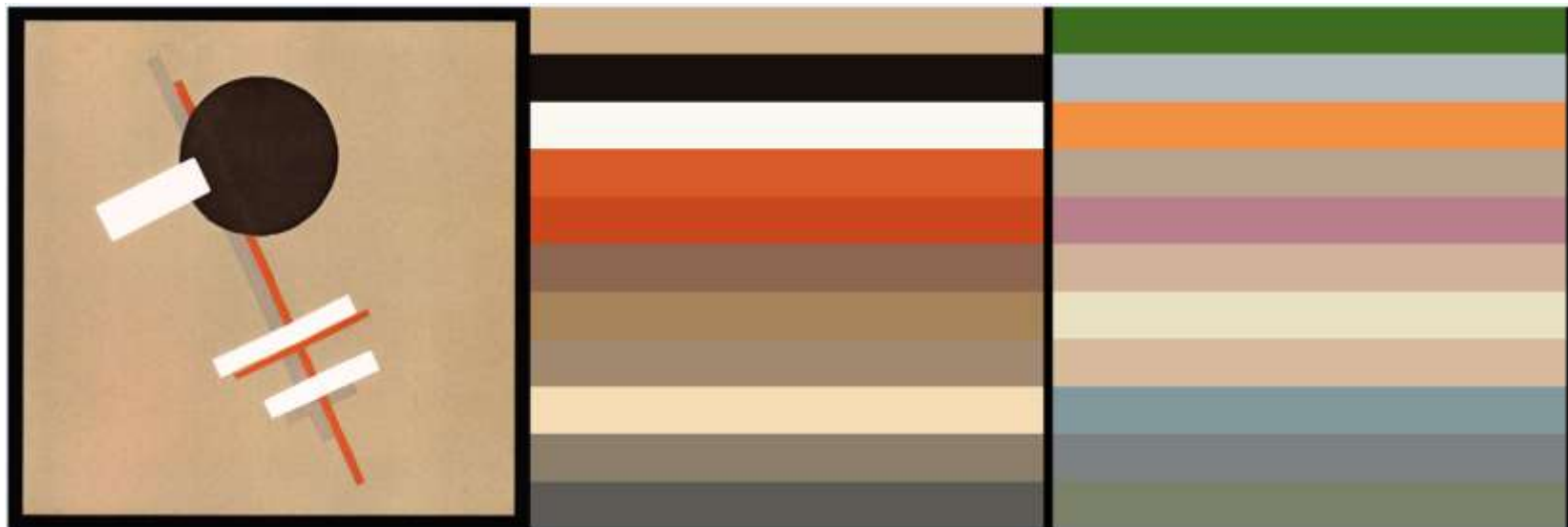
## Параметры выходной обложки:

1. Формат SVG.
2. Подпись музыкальной композиции на обложке



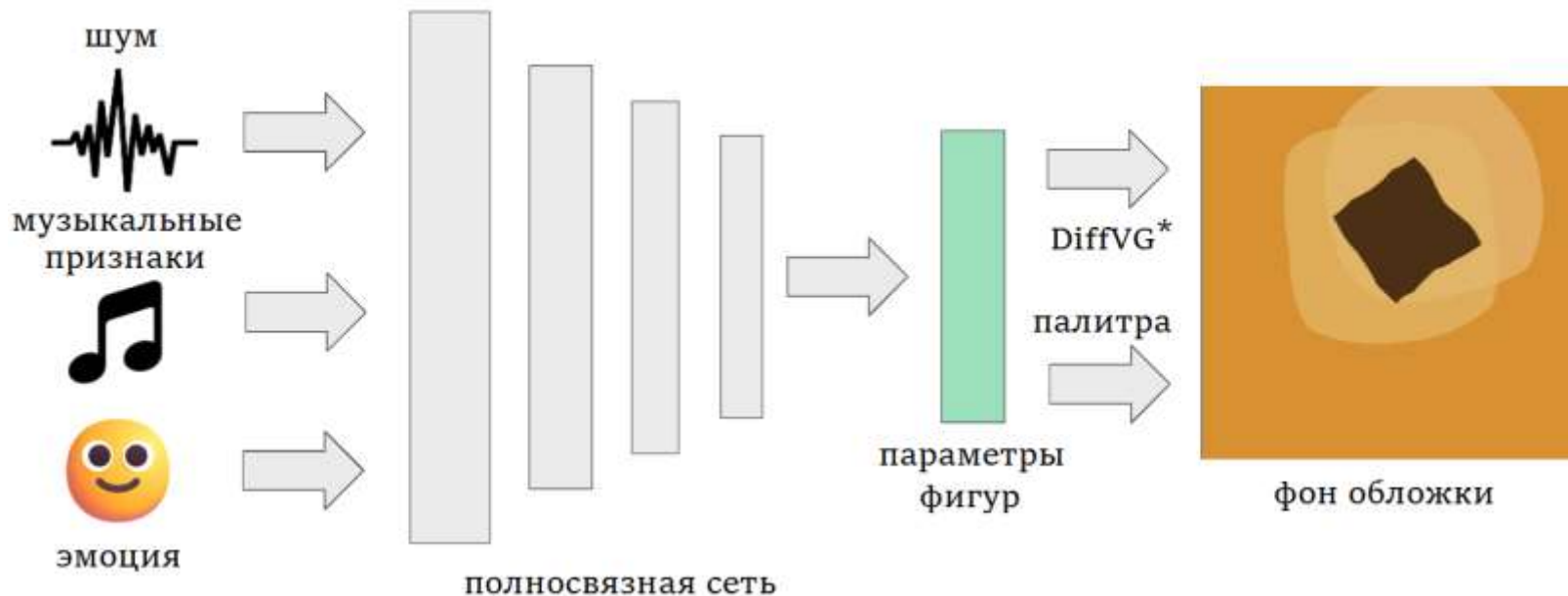
# Взаимосвязь музыки и цвета

Исходная обложка    Исходная палитра    Предсказанная палитра



Трек: «Кино – Группа Крови»

# Архитектура генератора CoverGAN



\* DiffVG – дифференцируемый растеризатор векторной графики

## Вставка подписей



### Параметры вставки:

- однородный фон;
- крупный размер шрифта;
- контрастный цвет текста;
- при необходимости: вертикальные буквы или перенос по словам.

# Примеры обложек CoverGAN





# Сравнение с другими моделями

DALL-E title



DALL-E lyrics



AttnGAN



AttnGAN + covers CoverGAN (2021) CoverGAN (2022)

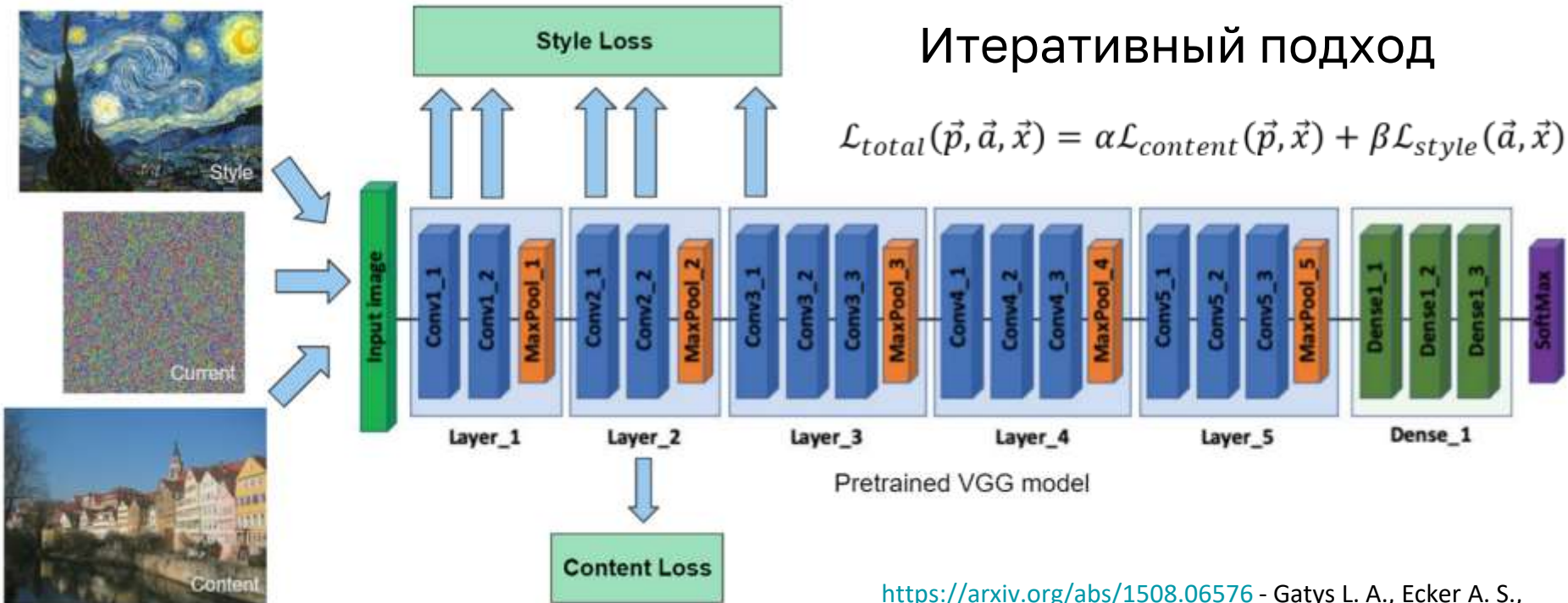


# Сравнение с другими моделями



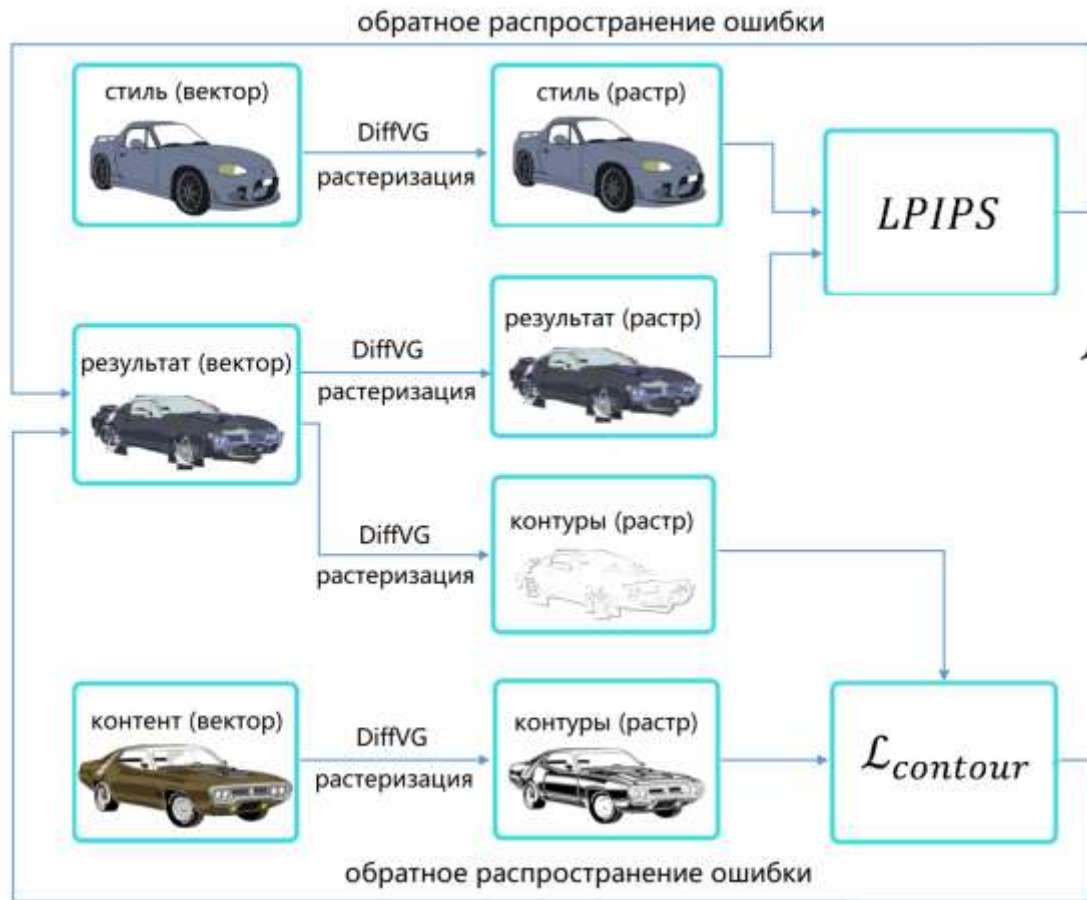
Метод	Общая оценка	Среди музыкантов
DALL-E title	0,34 ± 0,03	0,31 ± 0,05
DALL-E lyrics	0,30 ± 0,03	0,26 ± 0,05
AttnGAN	0,44 ± 0,04	0,23 ± 0,06
AttnGAN + covers	0,36 ± 0,03	0,19 ± 0,04
CoverGAN (2021)	0,68 ± 0,05	0,71 ± 0,05
<b>CoverGAN (2022)</b>	<b>0,74 ± 0,03</b>	<b>0,76 ± 0,04</b>

# VectorNST. Базовая архитектура



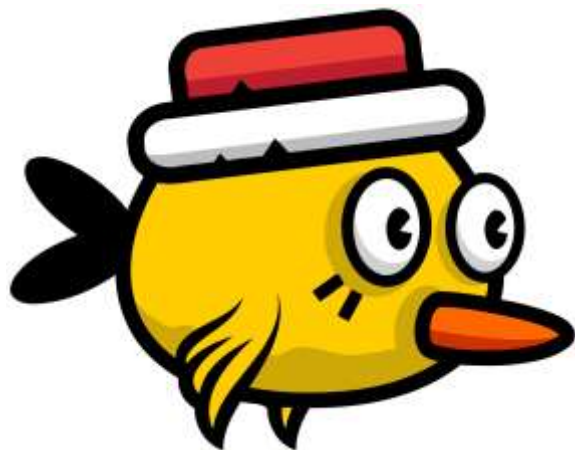
<https://arxiv.org/abs/1508.06576> - Gatys L. A., Ecker A. S., Bethge M. A neural algorithm of artistic style //arXiv preprint arXiv:1508.06576. – 2015.

# VectorNST

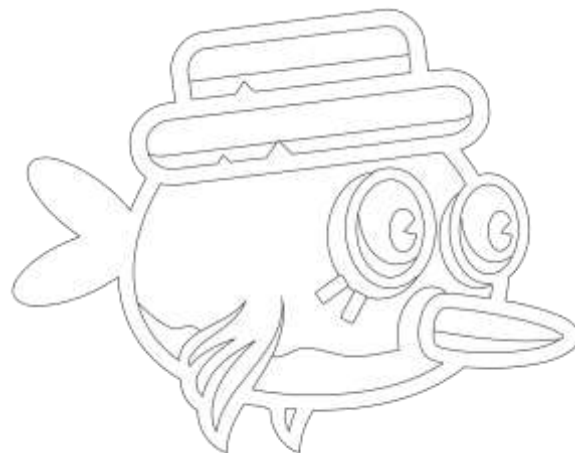


$$\mathcal{L} = LPIPS(x, y) + \lambda \mathcal{L}_{contour}(x, z)$$

Jarsky, I.; Efimova, V.; Chebykin, A.; Shalamov, V. and Filchenkov, A. (2024). Neural Style Transfer for Vector Graphics. In Proceedings of the 19th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 3: VISAPP; ISBN 978-989-758-679-8; ISSN 2184-4321, SciTePress, pages 686-693. DOI: 10.5220/0012438200003660. <https://www.scitepress.org/Papers/2024/12/4382/124382.pdf>



Входное векторное изображение







































Его контур

$$LPIPS(x, y) = \frac{1}{L} \sum_{l=0}^L \frac{1}{H_l W_l C_l} \sum_{h,w} \|\hat{x}_{hw}^l - \hat{y}_{hw}^l\|_2^2$$

$$\mathcal{L}_{contour}(x, z) = \frac{1}{n} \sum_{i=1}^n |x_i - z_i|$$



# Сравнение с другими моделями

Content Image	Style Image	Ours (vector)	DiffVG (vector)	Gatys <i>et al.</i> (raster)	StyTr <sup>2</sup> (raster)	SANet (raster)	CAST (raster)	ADP (raster)
								
								
								
								

# Сравнение с другими моделями



Время работы (в секундах)

Чем меньше – тем лучше

Маленькое: 256 × 256 пикселей, фигур < 100

Среднее: 512 × 512 пикселей, 100 ≤ фигур < 700

Крупное: 1024 × 1024 пикселей, 700 < фигур

Метод	Маленькое (сек)	Среднее (сек)	Крупное (сек)
VectorNST	5,93	33,52	112,10
DiffVG	4,20	26,21	98,57
Gatys et al.	1,61	4,14	11,59
StyTR <sup>2</sup>		0,3125	
SANet		0,0384	
CAST		0,0015	

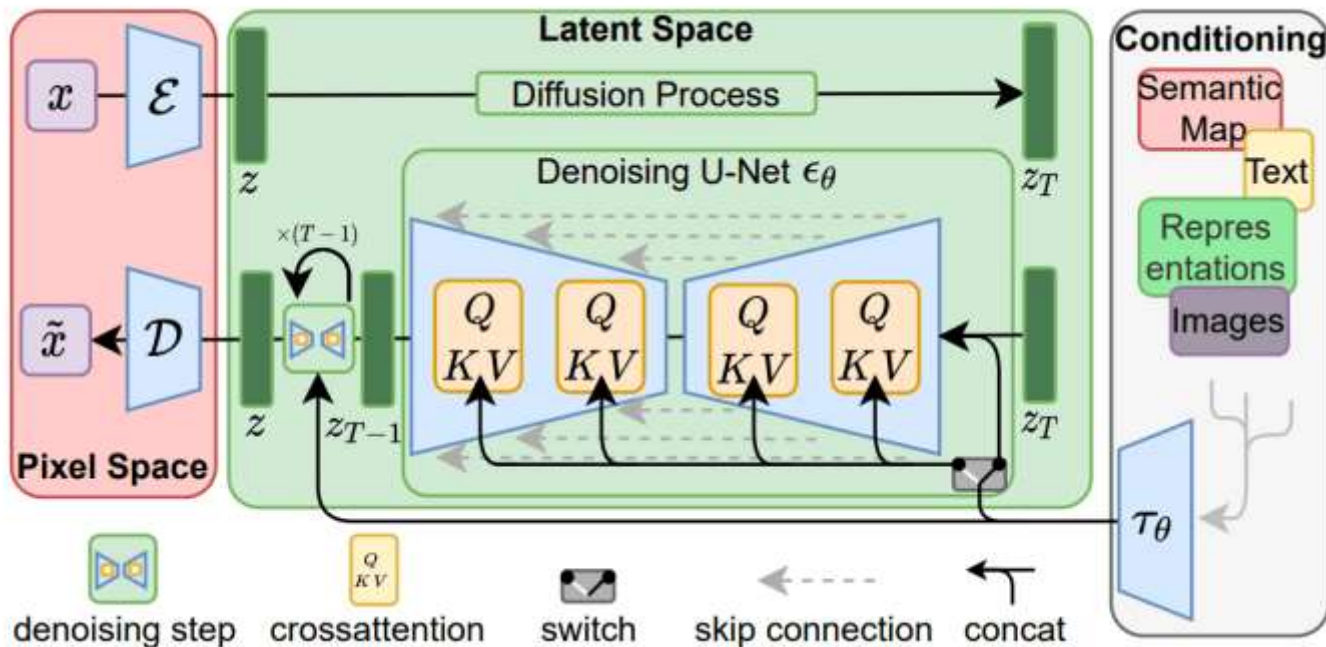
Результаты опроса

Чем больше – тем лучше

(метод доверит. интервалов)

Метод	Оценка
VectorNST	<b>0,56 ± 0,04</b>
DiffVG	0,44 ± 0,05
Gatys et al.	0,42 ± 0,06
StyTR <sup>2</sup>	<b>0,62 ± 0,05</b>
SANet	0,43 ± 0,06
CAST	<b>0,59 ± 0,06</b>
Attentioned Deep Paint	0,11 ± 0,04

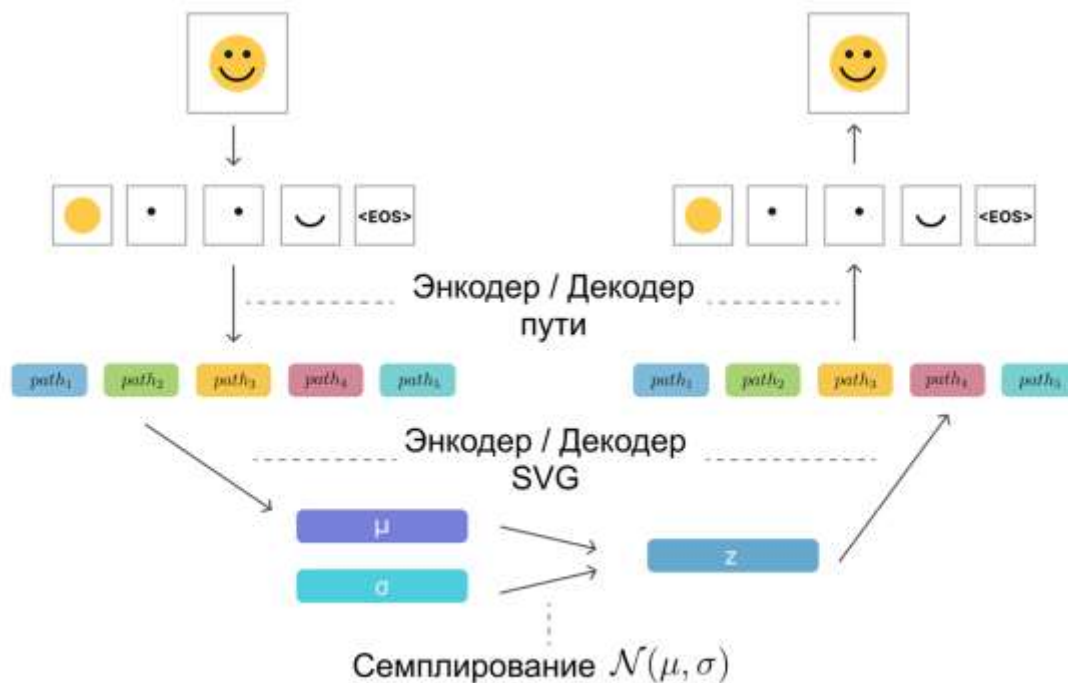
# VectorWeaver. Stable Diffusion



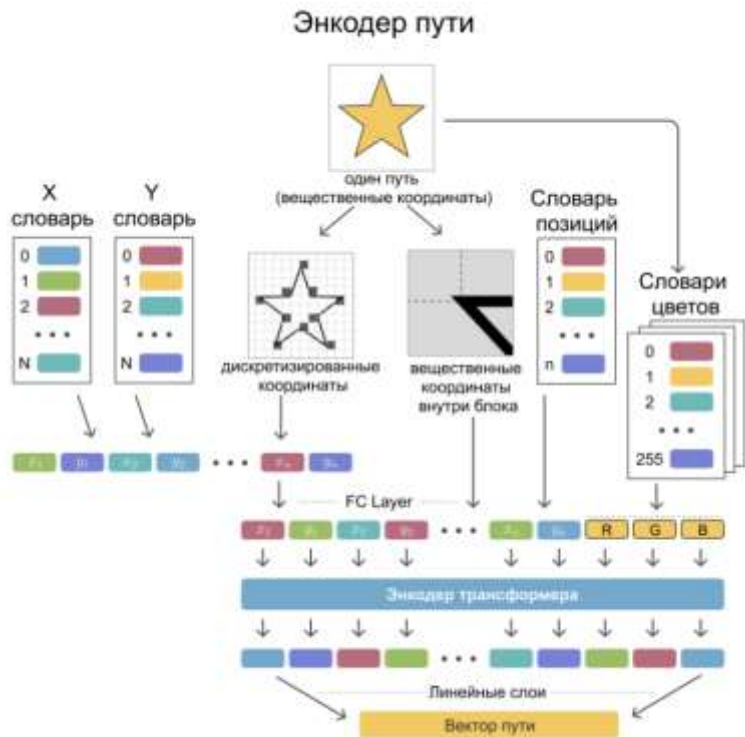
Можно выучить автоэнкодер и учить диффузную модель не на оригинальных изображениях, а в скрытом пространстве



# VectorWeaver. Автоэнкодер



# VectorWeaver. SVG Encoder



# VectorWeaver. SVG Decoder



# Результаты

# ІТМО



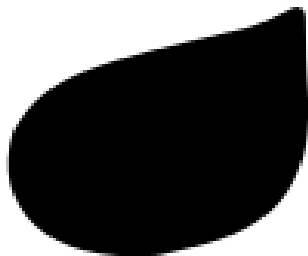
# Сравнение с конкурентами

Наш подход    DeepSVG    VectorFusion

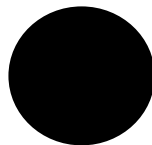


Метод	IS↑	FID↓	Оценка ассессора↑	Среднее время генерации, сек↓
DeepSVG	1.53	30.7	$0.19 \pm 0.03$	<b>0.06</b>
VectorFusion	1.2	15.1	$0.48 \pm 0.05$	1800
Наш подход	<b>1.8</b>	<b>4.3</b>	$0.79 \pm 0.03$	3

- Инициализация кривых кругами
- На основании детерминированного алгоритма SvgTracer (выбран он)



Исходное  
изображение

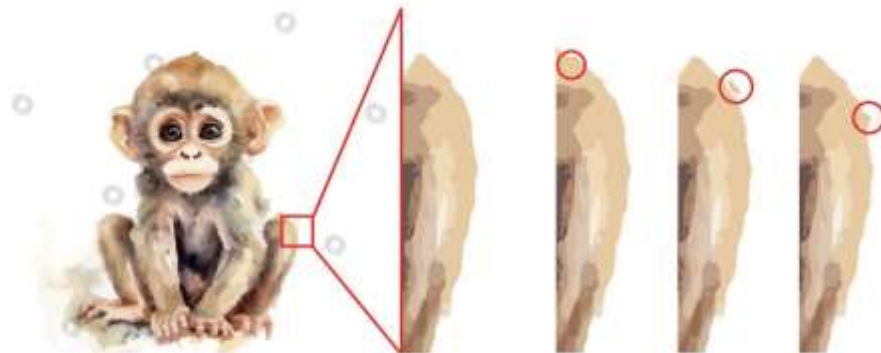


Инициализация  
кругами



SvgTracer

- Игольчатая мутация (вероятностное увеличение или уменьшение координат кривых)
- Мутация удаления путей или сегментов пути
- Мутация градиентного слияния путей

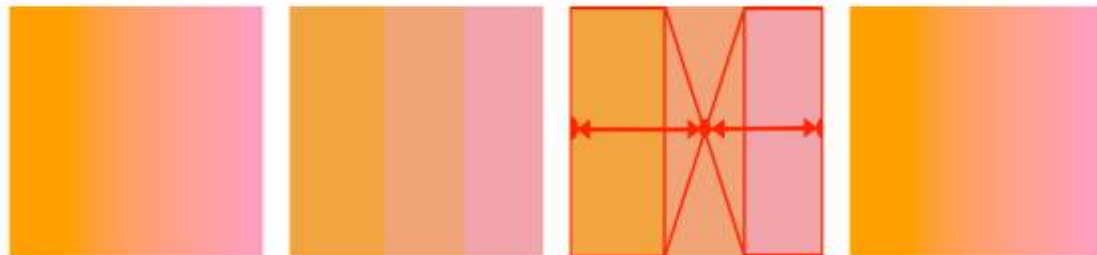


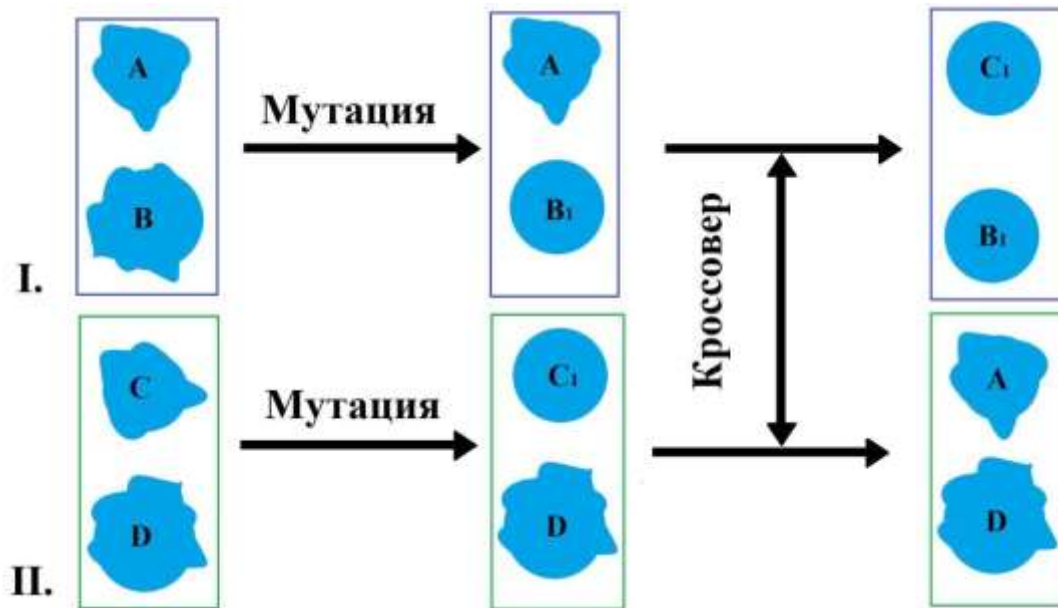
Исходное изображение

A

B

B





Вероятностный обмен путями



1. 
$$F_1(N, M) = \sum_{i=0}^w \sum_{j=0}^h |N_{i,j} - M_{i,j}|$$

2. 
$$F_2(N, M) = \sum_{i=0}^w \sum_{j=0}^h e^{|N_{i,j} - M_{i,j}| * \lambda + c}$$

3. 
$$F_3(N, M) = \sum_{i=0}^w \sum_{j=0}^h (|N_{i,j} - M_{i,j}| * \lambda)^2$$

В каждой из формул **N** и **M** - растровые изображения (RGB) размера **w** на **h**, где **w** - ширина изображения, а **h** - высота.

# EvoVec. Сравнительная таблица

Исходное изображение	LIVE (N=64)	DiffVG (N=512)	SvgTracer	Мой алгоритм	Метрика
					Результат
	51,5 м.	45,4 м.	7,5 с.	24,5 м.	Время работы
	4760	4255	1780	1547	Значение ф. отбора
	64	512	1790	985	Кол-во путей

# Заключение

- Представлен метод VectorMem позволяющий анализировать высокоразмерных векторные изображения и являющимся первым методом в данной области
- Разработан сервис CoverGAN который позволяет генерировать качественные векторные изображения (обложки) с использованием различных признаков
- Создана модель переноса стиля в векторной и растровой графике VectorNST на основе VGG модели и модифицированных функций ошибки
- Показан метод генерации векторных изображений в нативном формате VectorWeaver который позволяет генерировать простые векторные изображения качественнее альтернатив
- Разработан метод векторизации растрового изображения EvoVec который основан на эволюционном алгоритме и позволяет векторизовать изображение более качественно за меньшее время по сравнению с другими методами в данной области
- Ссылка на репозиторий: <https://github.com/CTLab-ITMO/VGLib>

**Спасибо  
за внимание!**

**iTMO** *re than a*  
**UNIVERSITY**