# 手書き線画の自動着色

九州大学芸術工学府芸術工学専攻博士2年 秋田 健太

- 私自身の研究について
  - 研究背景
  - 関連研究
  - 手法
  - 結果
- 昨今のイラストに関する研究について
  - Diffusion model登場以前
  - Diffusion model登場以降
  - 深層学習と著作権

- 私自身の研究について
  - 研究背景
  - 関連研究
  - 手法
  - 結果
- 昨今のイラストに関する研究について
  - Diffusion model登場以前
  - Diffusion model登場以降
  - 深層学習と著作権

- 私自身の研究について
  - 研究背景
  - 関連研究
  - 手法
  - 結果
- 昨今のイラストに関する研究について
  - Diffusion model登場以前
  - Diffusion model登場以降
  - 深層学習と著作権

#### 私自身の研究について

- 研究テーマは手書き線画の自動着色である。
- 通常、カラーイラスト制作は大まかに、ラフ→線画→カラーの 段階で製作され、線画→カラーを自動化することで、カラーイ ラスト制作を効率的に行えるようにすることを目指している。



#### イラストとスパコン

- 近年のコンテンツ制作支援の研究のほとんどは深層学習を利用 している。
- 解像度が要求される場面も多く、計算リソースが必要。 →スパコンの利用に繋がる。







解像度が低すぎると画像そのものの品質が低下してしまう。(左から128、256、512pixel)

© haruka

- 私自身の研究について
  - 研究背景
  - 関連研究
  - 手法
  - 比較·結果
- 昨今のイラストに関する研究について
  - Diffusion model登場以前
  - Diffusion model登場以降
  - 深層学習と著作権

#### 研究背景

- 深層学習の発展に伴い、線画の自動着色の研究も多く行われている。しかしながら、手書き線画に対応していないという大きな問題がある。
- 本研究では、手書き線画に対応することでより実践的な自動着 色を可能にする。

- 私自身の研究について
  - 研究背景
  - 関連研究
  - 手法
  - ・比較・結果
- 昨今のイラストに関する研究について
  - Diffusion model登場以前
  - Diffusion model登場以降
  - 深層学習と著作権

- 私自身の研究について
  - 研究背景
  - 関連研究
  - 手法
  - 比較·結果
- 昨今のイラストに関する研究について
  - Diffusion model登場以前
  - Diffusion model登場以降
  - 深層学習と著作権

#### 手法

• 本手法では、Semantic segmentation mask(領域が意味的に、例えば髪や肌、分割されている画像)を用い、参照画像のテクスチャ詳細を線画に転写する。



カラー画像



Segmentaion mask

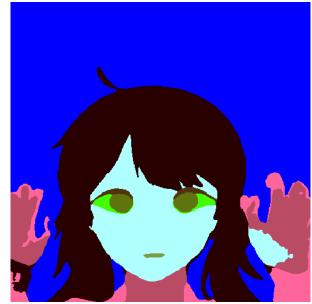
### データセットの作成

- 本手法では従来通りカラー画像から作成する線画に加えて Segmentation maskを作成する。
- データセットは3種類あり、カラー画像( $I_{color}$ )、カラー画像から作成した線画( $I_{line}$ )、Segmentation mask( $I_{seg}$ )である。





線画



Segmentation mask

© yuichi

# Segmentation maskの作成

- Segmentation maskはFew-shot手法[Tritrong+ '21]により作成。
- 使用するFew-shot手法ではStyleGAN2を用いるため、カラー画像のデータセットで事前学習する。

# Segmentation maskの作成

- Segmentation maskを作成するためにSQUIDを使用している。
- StyleGAN2は転移学習しており、GPU並列で学習した。
- 潜在変数への変換は、各GPUで独立で行った。

	GPU	時間
StyleGAN2の学習	SQUID・GPUノード	約48時間
画像の潜在変数への変換	SQUID・GPUノード	数十時間×8ノード

#### 線画の作成

- 既存着色手法がテクスチャ詳細を表現できない理由の一つにドメインギャップがある。
- このドメインギャップを小さくするために学習ベースの線画抽出法を用いる。



カラー画像



通常の抽出線画

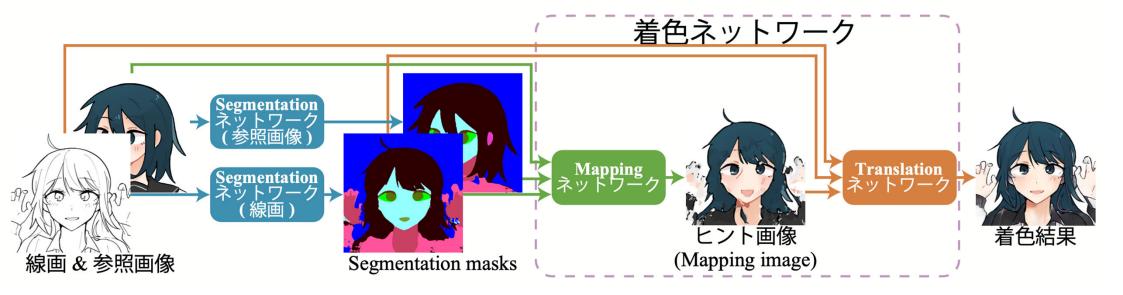


学習ベースの抽出線画

© yuichi

## 線画の着色システム

テスト時は、線画と参照画像のsegmentationを行い、それを元に着色を行う。着色ネットワークはCoCosNet v2 [Zhou+ '21]を用いた。



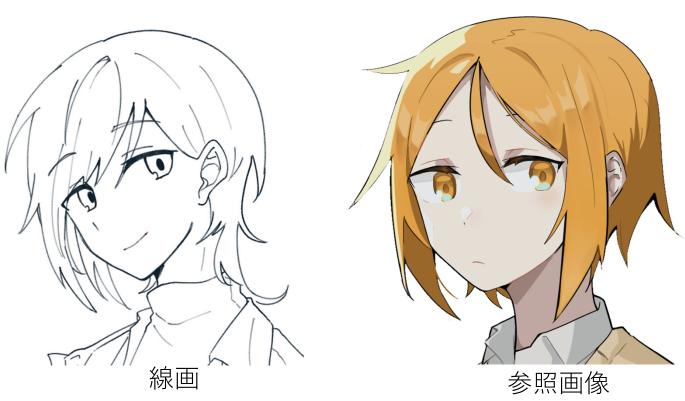
#### 着色ネットワークの学習

- 着色ネットワークでもSQUIDを使用している。
- 着色ネットワークの学習は40時間ほどかかる。
- 研究室の個人用パソコンで3090を2枚搭載しているが、そのパソコンでは2週間弱かかる。

- 私自身の研究について
  - 研究背景
  - 関連研究
  - 手法
  - 結果
- 昨今のイラストに関する研究について
  - Diffusion model登場以前
  - Diffusion model登場以降
  - 深層学習と著作権

# 結果

本手法は参照画像のテクスチャの詳細を線画に転写した着色ができている。





着色画像

- 私自身の研究について
  - 研究背景
  - 関連研究
  - 手法
  - 結果
- 昨今のイラストに関する研究について
  - Diffusion model登場以前
  - Diffusion model登場以降
  - 深層学習と著作権

- 私自身の研究について
  - 研究背景
  - 関連研究
  - 手法
  - 結果
- 昨今のイラストに関する研究について
  - Diffusion model登場以前
  - Diffusion model登場以降
  - 深層学習と著作権

#### 昨今のイラストに関する研究について

- 私が研究しているのは主に自動着色についてであるが、イラストやそれに類するコンテンツ制作に関する研究は他にも多く行われている。
- ここでは基本的に深層学習を用いた研究を取り上げる。
- 特に昨年からDiffusion modelによる生成モデルが話題になっている(研究だけでなく開発的側面も強い)。

#### そもそもDiffusion modelとは

- 学習画像(右の画像)にノイズを足していき、ノイズ画像を作成(左の画像)する。その後逆にノイズから学習画像を復元する。学習についてはノイズ除去についてのみ行う。
- 今主流の方法は潜在変数でこの処理を行う(Stable Diffusionと呼ばれる)。
- 処理は重いが、Generative Adversarial Nets (GAN)より学習が 安定している。

- 私自身の研究について
  - 研究背景
  - 関連研究
  - 手法
  - 結果
- 昨今のイラストに関する研究について
  - Diffusion model登場以前
  - Diffusion model登場以降
  - 深層学習と著作権

#### Diffusion model登場以前

- GANを使った研究が多く行われている。
- コンテンツ制作を支援する研究、完成したコンテンツを編集する研究、コンテンツ自体を生成する研究等がある。

#### Diffusion model登場以前

- ほとんどの研究はある一定の結果は示しているものの、研究レベルに留まっていた→精度の問題だけでなくモデルが提供されない・されてもある程度の知識がなければ使えない。
- ・データを自作していることが多く、そのデータが提供されない。→基本的に第三者の著作物を利用しているため? (法的には問題ないが、気持ち的な問題?)

- 私自身の研究について
  - 研究背景
  - 関連研究
  - 手法
  - 結果
- 昨今のイラストに関する研究について
  - Diffusion model登場以前
  - Diffusion model登場以降
  - 深層学習と著作権

#### Diffusion model登場以降

- これまでは研究レベルで注目される研究が多かった。しかしながら、一般の人々にまで広まることはほとんどなかった。
- Diffusion model登場以降、特に昨年のNovelAIの登場から一般でも注目されるようになってきた。
- 研究だけでなく、アプリ等の開発が精力的に行われている。



Diffusion modelにより生成した画像(モデルはAnything v3.0)

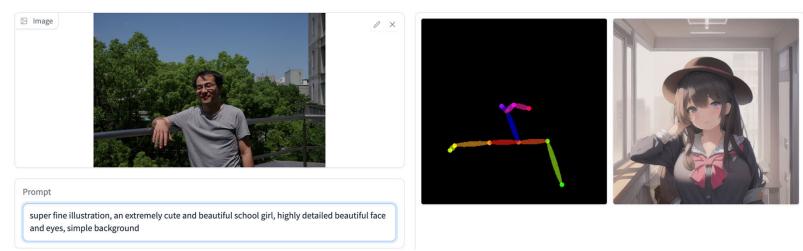
#### アプリケーションの開発

- Hugging Faceを利用したものが多く、どんな人でも簡単に試す ことができる。
- Colaboratoryを使っている場合もあり、情報が多く簡単に利用できると考えられる。

#### 技術の発展速度

- 研究的な観点で見れば、既存の研究に対してどのように Diffusion modelを適用するかという部分が大きい。ただし、できる幅が急激に広がった。
- 一方で、開発は非常に早く、発表された論文が数日で一般の人でも利用できるレベルになっていたりする。

#### **Control Stable Diffusion with Human Pose**



Hugging Faceで公開されているアプリで生成した結果

- 私自身の研究について
  - 研究背景
  - 関連研究
  - 手法
  - 結果
- 昨今のイラストに関する研究について
  - Diffusion model登場以前
  - Diffusion model登場以降
  - 深層学習と著作権

#### 深層学習と著作権

- ・深層学習ではデータを扱うため、そのデータが第三者のものであれば第三者が著作権を持っている。
- 学習に限っていえば、無断で第三者の著作物を学習に使うことは合法的であり、ほとんどの場合気にする必要はない。
- ただし、著作者の利益を害する場合はその限りではない。 ("AI イラストや"が話題になったかと思うが、これを無断で行った場合違法であると考えられる。)

#### 深層学習と著作権

- 学習の場合は多くの場合大きな問題はない。ただし、生成物については異なる。
- Diffusion modelは精巧な画像を生成できるため、これまでより 著作権についての議論を見かけるようになった。
- 結論から言えば、画像を生成した者が意図的に著作権侵害になりうる画像を生成する・明示的に特定作者のような画像であるということを主張しない限り問題ないとされている。

#### 深層学習と著作権

- これまでの話はあくまで法律的な話である。
- 感情的な話で言えば、自身の著作物を学習に使用されること、 またAIを用いて画像を生成することに抵抗があるイラストレー タは一定数いる。
- 逆に自身の苦手な作業をAIに代用して欲しい人も一定数おり、 支援のためのバランスが重要である。

#### 最後に

- ・深層学習はまだ発展途上の分野であるが、現時点でも十分な性能がありイラスト分野に影響を与えつつある。
- イラストにおける研究は第三者のデータを用いるため、著作権、またイラストレータの心象に配慮しつつ研究を進める必要がある。
- イラスト・アニメ制作支援の研究はこれまでのワークフローを 大幅に効率化することも期待でき、個人でも大規模な制作がで きることも期待できる。