

GANを用いた漫画の画像補完

横田 広之¹ 古田 諒佑² 谷口 行信¹ 日並 遼太³ 石渡 祥之佑³
 東京理科大学大学院 工学研究科¹ 東京大学 生産技術研究所² Mantra 株式会社³

1 はじめに

近年、海外における漫画の市場規模が拡大しており、日本語で書かれた漫画を別の言語に翻訳する需要が高まっている。漫画の文字には、吹き出しの中に書かれたものの他にイラストの上に直接書かれたものがあるため、翻訳した文字を書き込む前にあらかじめ元の文字領域を補完しておく必要がある。本研究では、漫画のイラストの上に書かれた文字を削除し、その領域を補完する手法を検討する。具体的には、漫画のイラスト部分の画像と文字マスクのペアを入力として、マスク領域を補完した画像を出力するGANのモデルを学習する。

2 関連研究

2.1 Navier-Stokes (NS) 法 [1]

Navier-Stokes 法は、流体力学に基づく偏微分方程式を用いたルールベースの画像補完手法である。具体的には、非欠損領域から欠損領域にかけて連続的に等輝度線を伝播することで欠損領域を補完する。

2.2 GLCIC [2]

GLCIC は、Generative Adversarial Network (GAN) を用いた画像補完手法である。画像全体の整合性を考慮する Global Discriminator と欠損領域のみの整合性を考慮する Local Discriminator を用いることで、より自然な画像補完を学習することができる。GLCIC の学習は3つのフェーズに分かれており、フェーズ1では Generator のみを学習し、フェーズ2では Generator の重みを凍結して Discriminator のみを学習し、フェーズ3では Generator と Discriminator の両方を学習する。

3 実験

漫画における画像補完の手法検討のため、漫画の画像と文字のアノテーションが含まれる Manga109 [3] データセットと文字のマスク、GAN のモデルである GLCIC を用いた予備実験を行った。また、ルールベースの画像補完手法である Navier-Stokes 法を実験の比較対象とし、ピーク信号対雑音比 (PSNR) を用いて精度評価を行った。GLCIC のパラメータは、入力サイズを 320×320 、学習ステップ数をフェーズ1で90000ステップ、フェーズ2で10000ステップ、フェーズ3で400000ステップとして学習を行った。また、Manga109 データセットのうち、ARMS, LoveHina_vol01, LoveHina_vol14 の3巻のデータから、1242枚の文字が含まれない画像を学習用80%、評価用20%に分割して使用した。

表1 Navier-Stokes 法と GLCIC の評価結果

	Navier-Stokes [1]	GLCIC [2]
(full) PSNR [dB]	28.72	30.20
(mask) PSNR [dB]	11.48	12.96

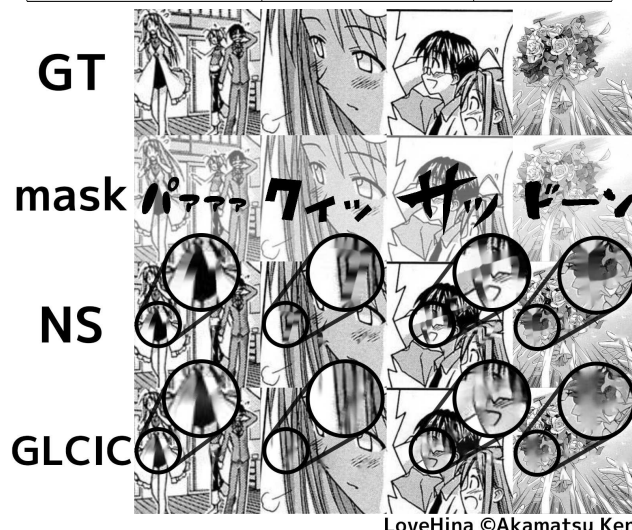


図1 Navier-Stokes 法と GLCIC の出力結果の例

4 実験結果

評価用データにおける Navier-Stokes 法と GLCIC の評価結果を表1に示す。PSNRにより精度評価を行った結果、Navier-Stokes 法よりも GLCIC の方が良い結果となった。次に、Navier-Stokes 法と GLCIC の出力結果の例を図1に示す。この結果から、マスクの領域に線画が多く含まれる場合、GLCIC を用いても出力結果がぼやけてしまうことが分かる。

5 今後の課題

漫画の画像補完における精度向上のため、線画とスクリーントーンを分けて補完することで出力結果を改善する手法を検討する。

参考文献

- [1] M. Bertalmio, et al., “Navier-Stokes, Fluid Dynamics, and Image and Video Inpainting,” in CVPR, 2001.
- [2] S. Iizuka, et al., “Globally and Locally Consistent Image Completion,” in SIGGRAPH, 2017.
- [3] K. Aizawa, et al., “Building a Manga Dataset “Manga109” with Annotations for Multimedia Applications,” IEEE MultiMedia, 2020.