

Sketch-based Food Image Generation and Editing

Cho Jaehyeong, 下田 和, 柳井 啓司(電気通信大学)

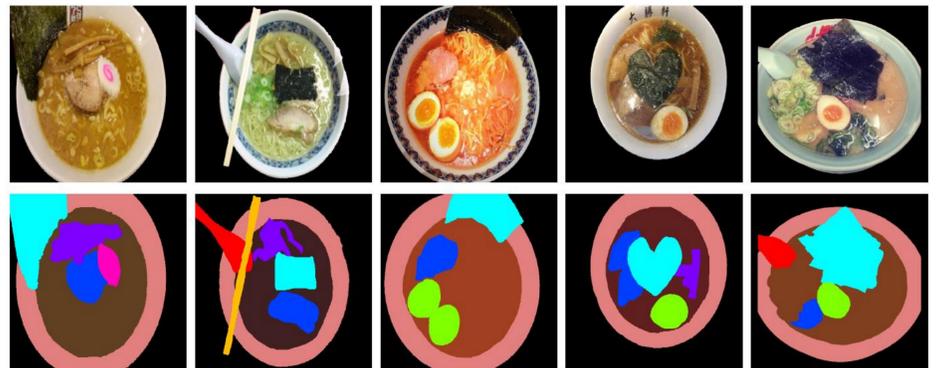
はじめに

- 画像を合成・修正は熟練したスキルと多くな時間が必要なタスク
→ 深層学習を用いて簡単に、そして、早く画像を修正
- Mask sketch画像を基づいて画像生成
- 実画像から生成されたsegmentation maskを基づいて画像を修正



データセット

- ラーメンの実画像とmask画像に構成されたデータセットを作成
- Mask画像は15classのlabelを含む
- 500枚をモデルの学習に使用



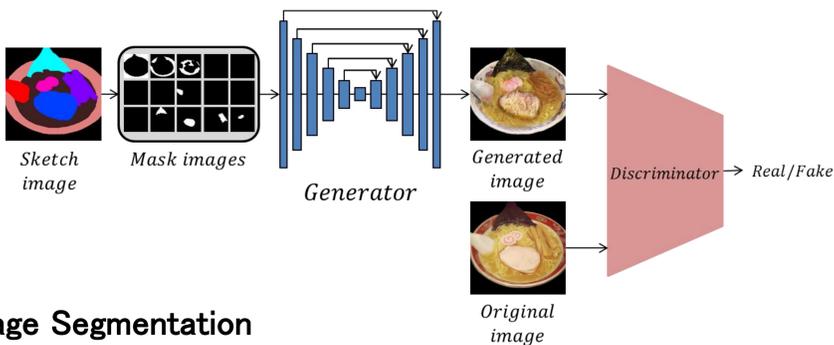
手法

- Image Generation
 - Pix2Pix [2]
 - Pair 画像間の違いを学習して、その差を補う形で画像を出力
 - Skip connectionを持つEncoder - Decoderを持つU-Net^[4]で構成
 - Generatorにスケッチした画像を入力
 - Discriminator に損失関数そのものを学習

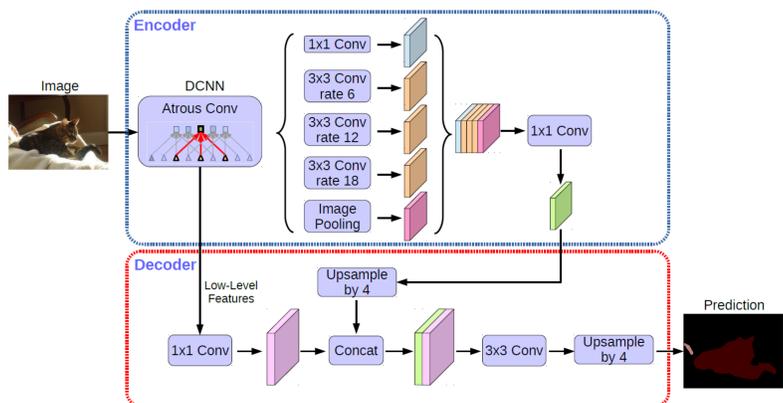
$$\mathcal{L}_{cGAN}(G, D) = \mathbb{E}_{x,y}[\log D(x,y)] + \mathbb{E}_{x,z}[\log(1 - D(x, G(x,z)))]$$

$$\mathcal{L}_{L1}(G) = \mathbb{E}_{x,y,z}[\|y - G(x,z)\|_1]$$

$$G^* = \arg \min_G \max_D \mathcal{L}_{cGAN}(G, D) + \lambda \mathcal{L}_{L1}(G)$$



- Image Segmentation
 - Deeplab V3+ [3]
 - 強力なencoderとシンプルで効果的なdecoder構造を持つ semantic segmentation model
 - 入力された画像の各要素をsegmentationしたmask画像を出力

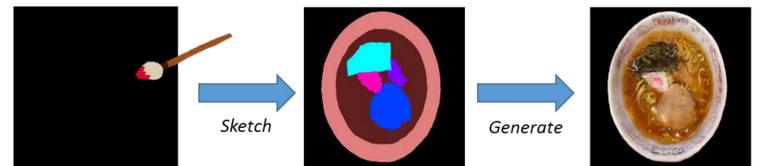


今後の課題

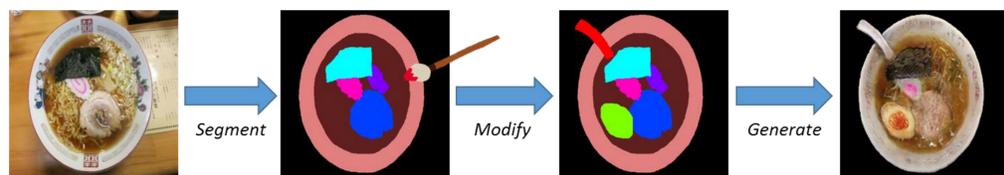
- 形状とスタイルの分離
- より簡単に画像修正できるインターフェースに改良
- 生成画像の品質の向上

実験

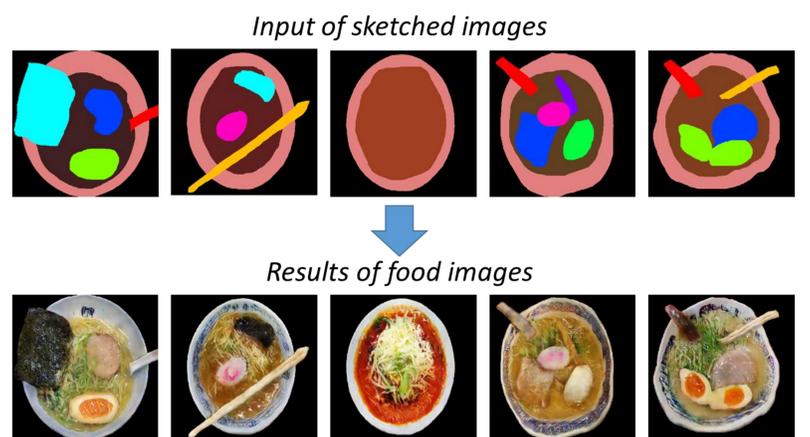
- 実験(1) スケッチ画像からラーメン画像生成



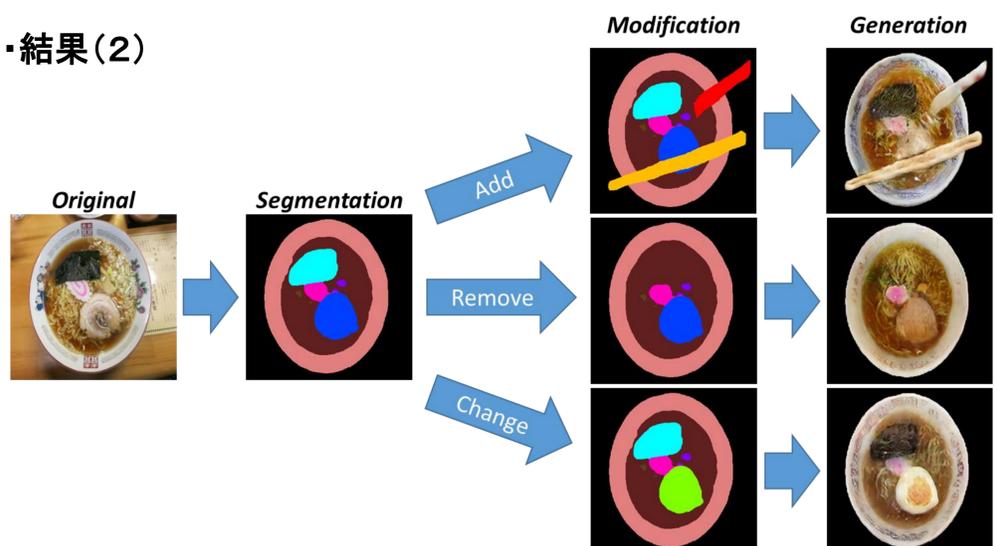
- 実験(2) 画像の領域分割と生成



- 結果(1)



- 結果(2)



参考文献

[1] Goodfellow, Ian, et al. "Generative adversarial nets." *NIPS*. 2014.
 [2] Isola, Phillip, et al. "Image-to-image translation with conditional adversarial networks." *CVPR*. 2017.
 [3] Chen, Liang-Chieh, et al. "Encoder-decoder with atrous separable convolution for semantic image segmentation." *ECCV*. 2018.
 [4] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." *MICCAI*. 2015.