

# Neural Style Transferと 領域分割による画像の部分的質感操作

松尾 真, 下田 和, 柳井 啓司

電気通信大学 大学院情報理工学  
研究科 総合情報学専攻

# 画像のスタイルの変換

- 「A Neural Algorithm of Artistic Style」
  - Gatys et al. Year 2015, arXiv:1508.06576
  - Deep Neural Networkを用いて、画像を絵画風に変換する
  - 元画像の構成を損なわずにスタイルを変換



コンテンツ画像



スタイル画像



# 画像内物体の質感操作への応用

- Flickr material dataset (FMD)
  - 10種類の素材画像(布、金属、木製など)
- 素材画像のstyle transfer
  - 皮を石に
  - ガラスを鉄に
- 物体の任意質感変換を目指す



FMD画像の例

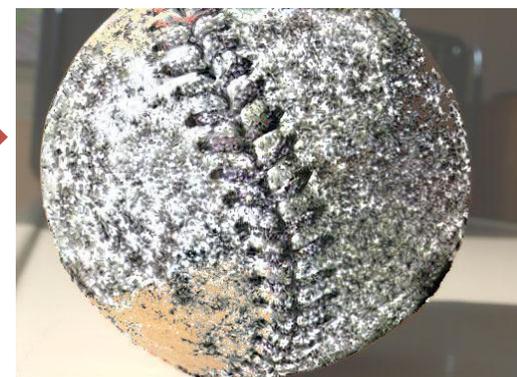
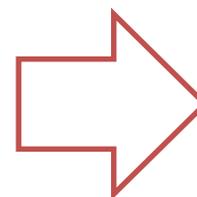
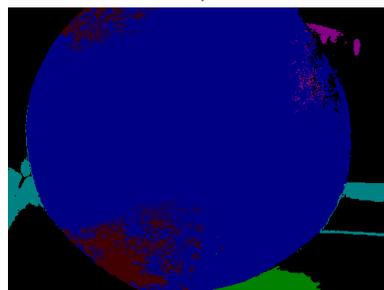
# スタイル変換＋領域分割

- Neural Style Transferは背景も変化させてしまう。
- 素材画像の領域分割と併用
  - 元画像の物体領域のみを変換

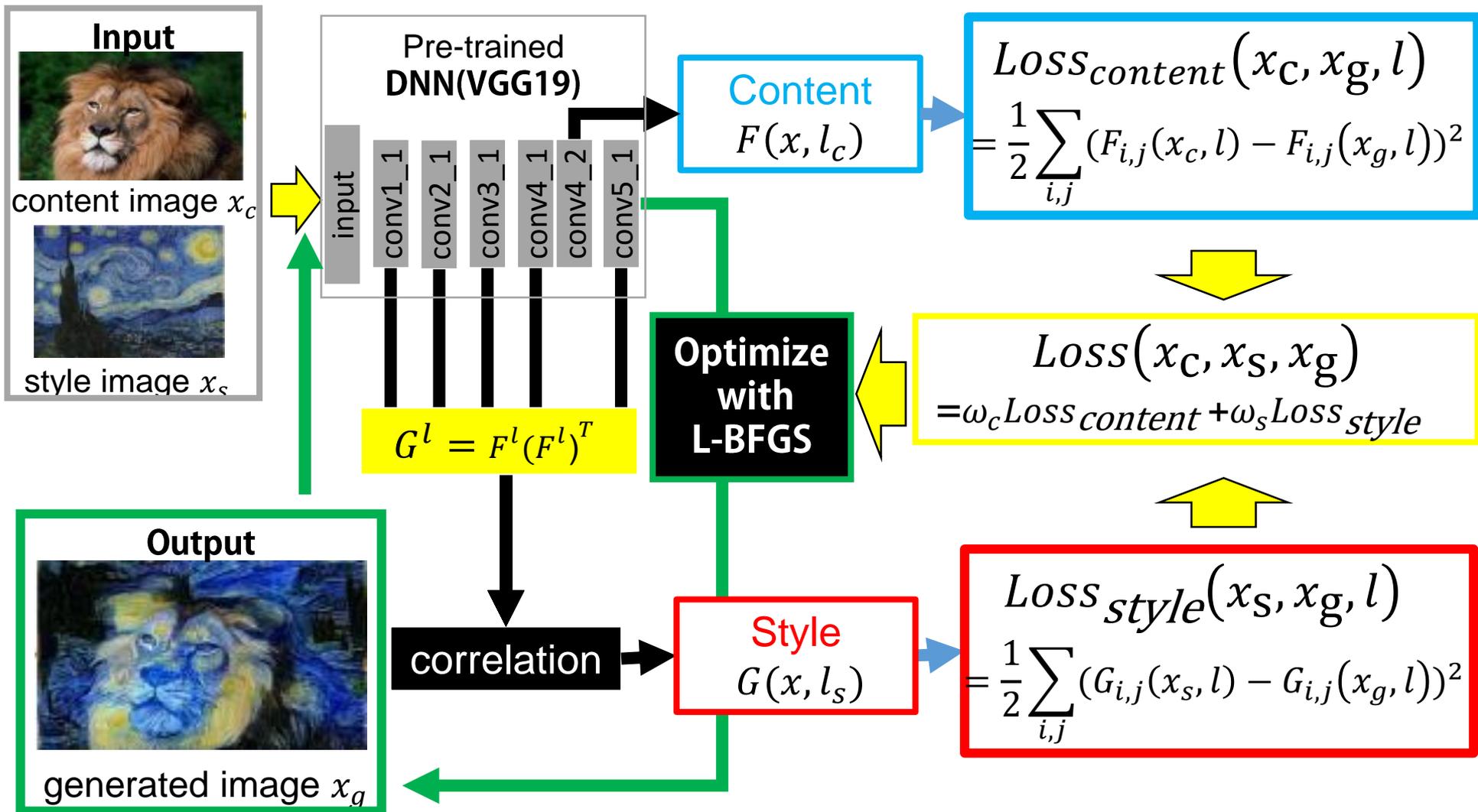


領域分割

スタイル変換



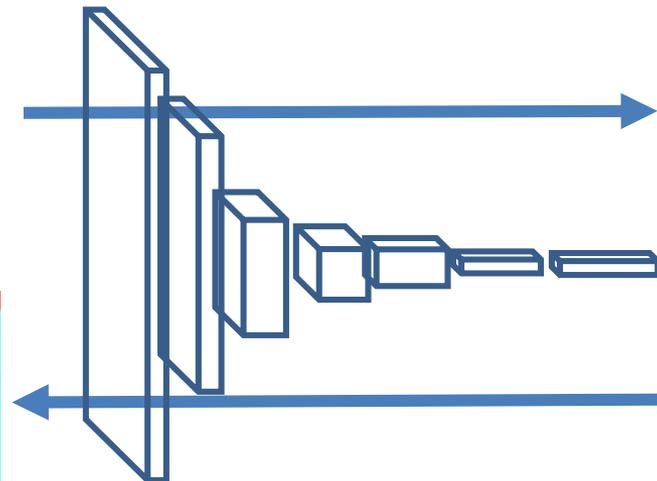
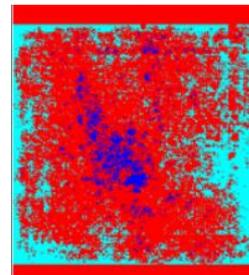
# スタイル変換の仕組み



# 領域分割

- 弱教師あり領域分割
  - Backpropagationを用いた伝搬値による位置推定

$$w = \left. \frac{\partial S_c}{\partial I} \right|_{I_0}$$



<認識結果>  
バッタ

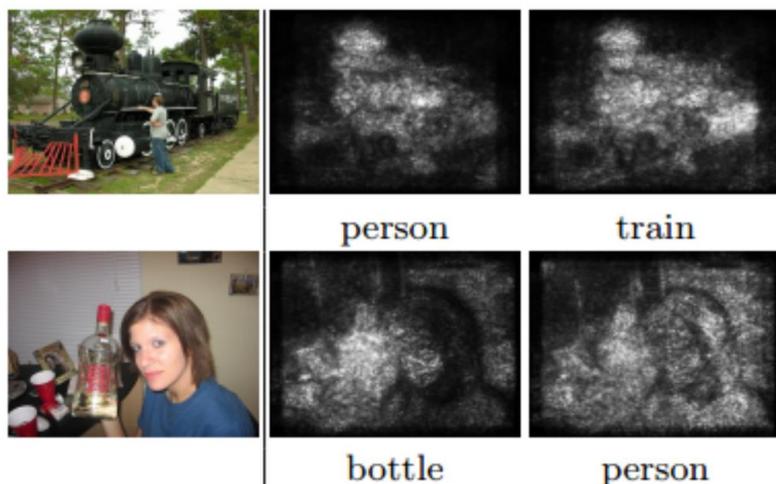
学習しようとする領域が反応  
これを活用し、位置を推定

Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps

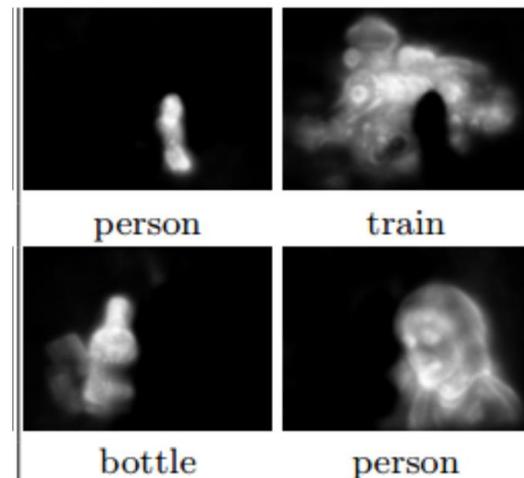
Simonyan et al.  
ICLR workshop 2014

# 弱教師有り領域分割

- 同研究室下田による改良案
  - より鮮明な、カテゴリ別の領域分割に改良
- MIRUで発表済み, ECCVで発表予定
- 将来的にWeb画像に応用
  - FMDのマスキは使用しない



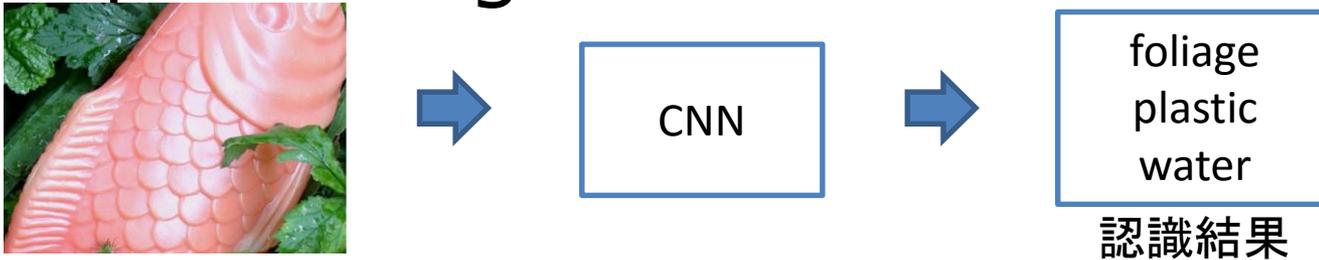
Simonyan et al. 国立大学法人 電気通信大学



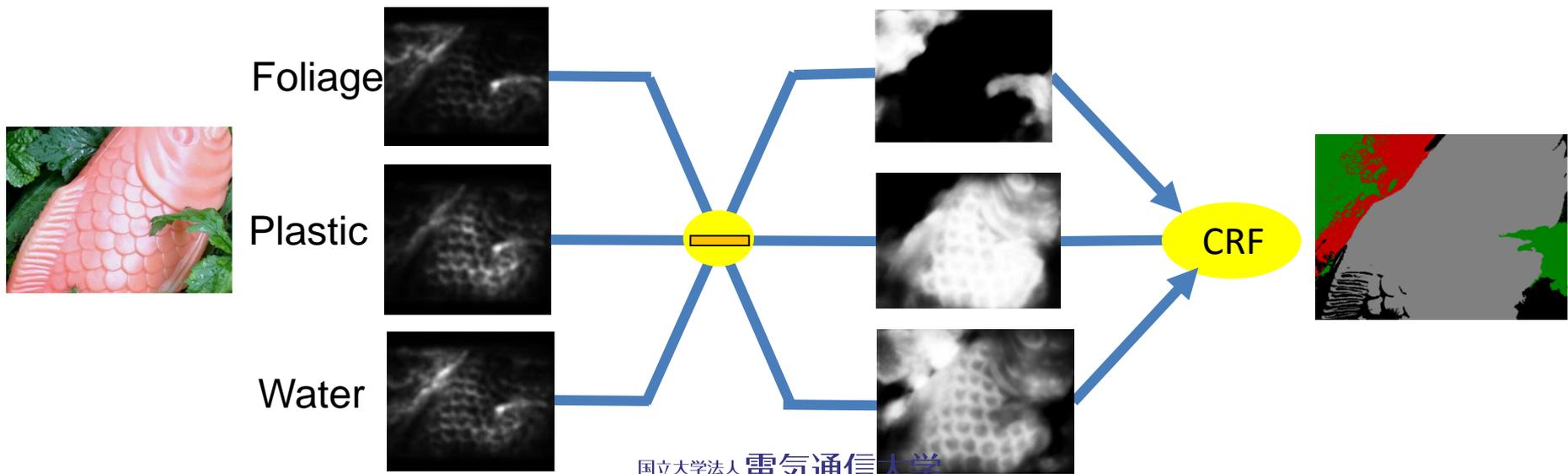
Ours

# 質感画像の弱教師あり領域分割

- Deep learningによる質感画像の認識



- 各カテゴリの認識に寄与した領域の推定



# 実験

## 使用したFMDの画像

- コンテンツ画像2枚(ball, jacket)
- スタイル画像200枚(各素材20枚)



全ての組み合わせで部分的質感変換を行い、  
変換後の領域分割の精度を評価

- Pixel accuracy, Mean IUを評価
- 正解はFMDのマスキ画像

# 実験の流れ

Materials



(1) Original images



(2) Style Transfer to whole regions



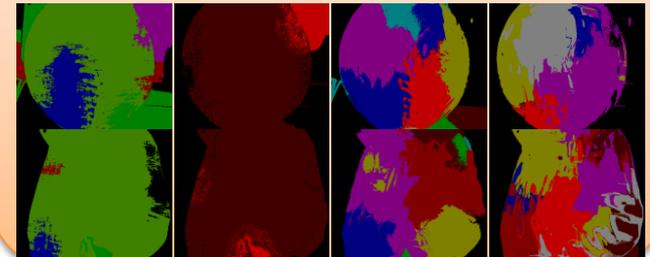
(3) Segmentation to original images



(4) Style synthesis to only the segmented regions

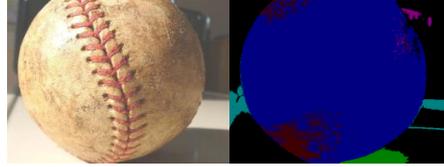


(5) Segmentation to the synthesized images

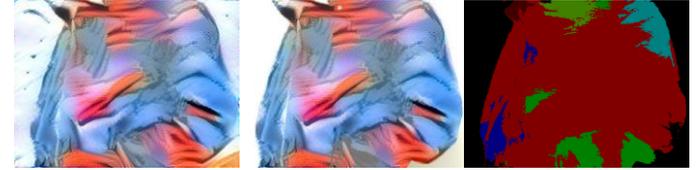
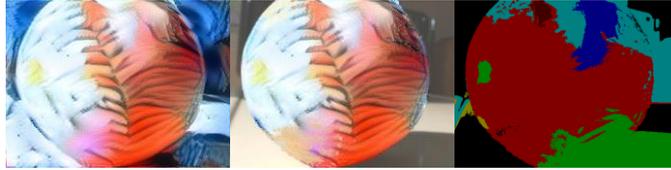


# 変換結果例

style content



fabric



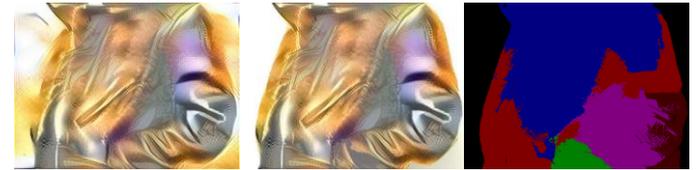
foliage



glass



leather

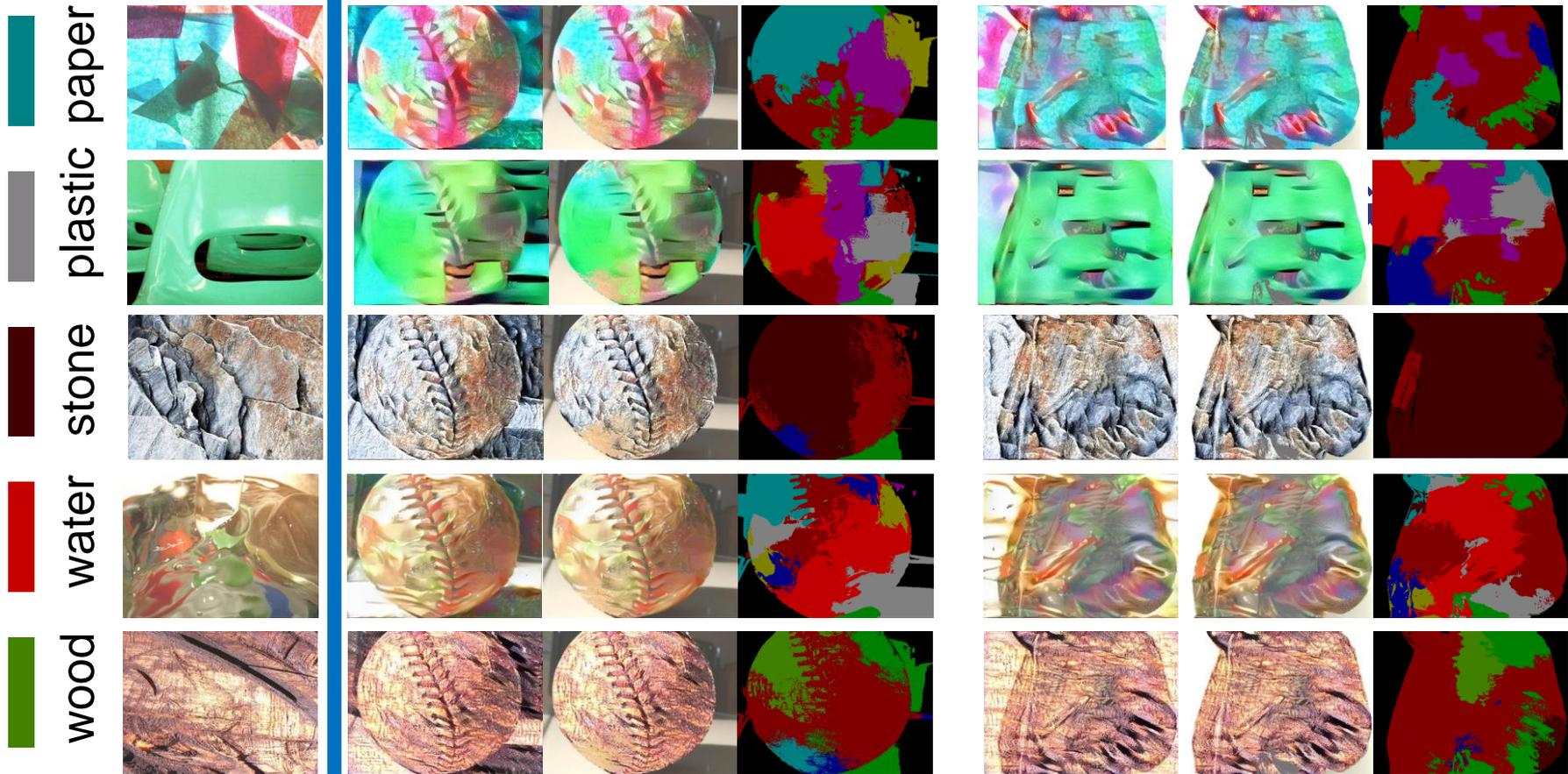
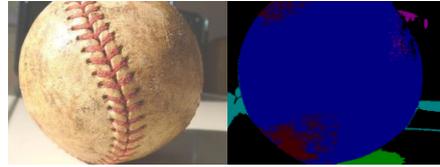


metal



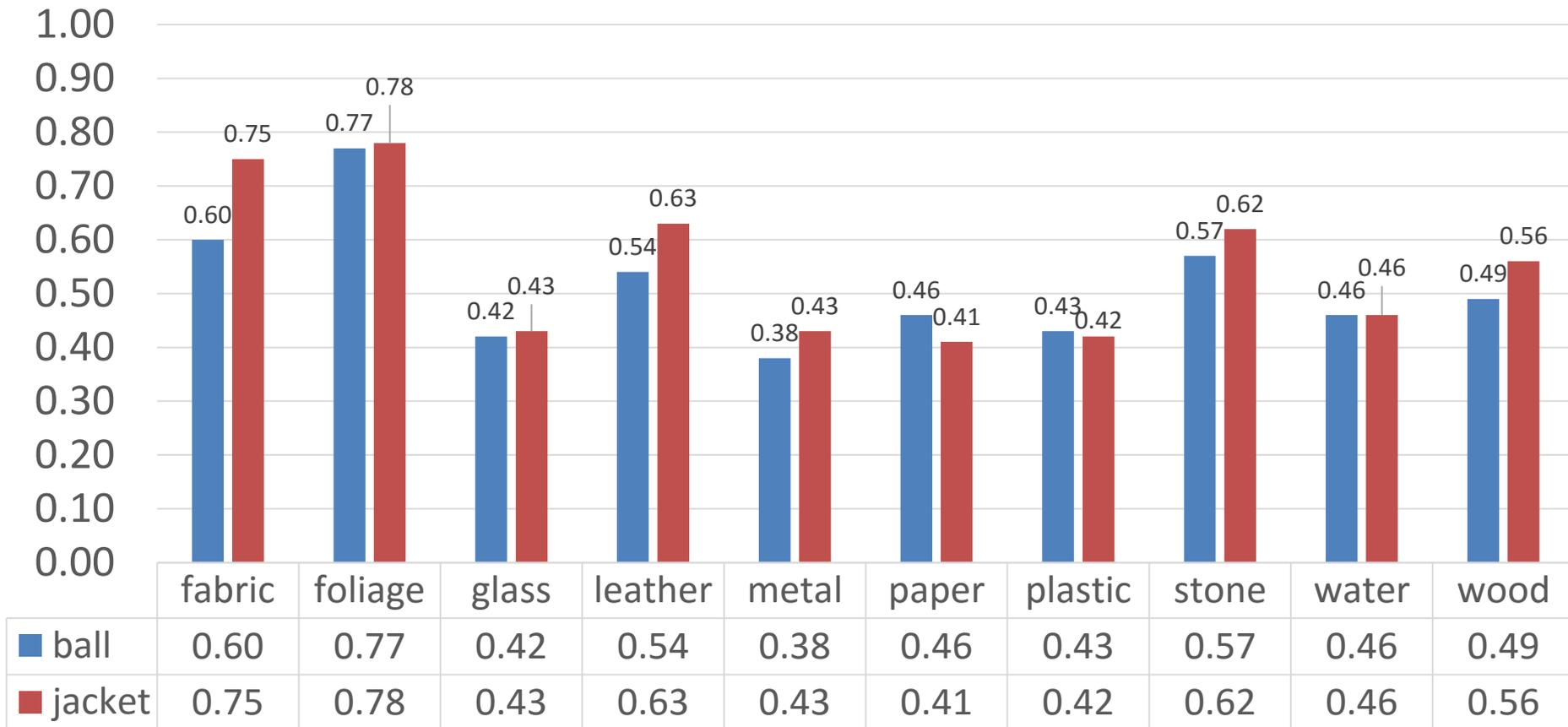
# 変換結果例

style content



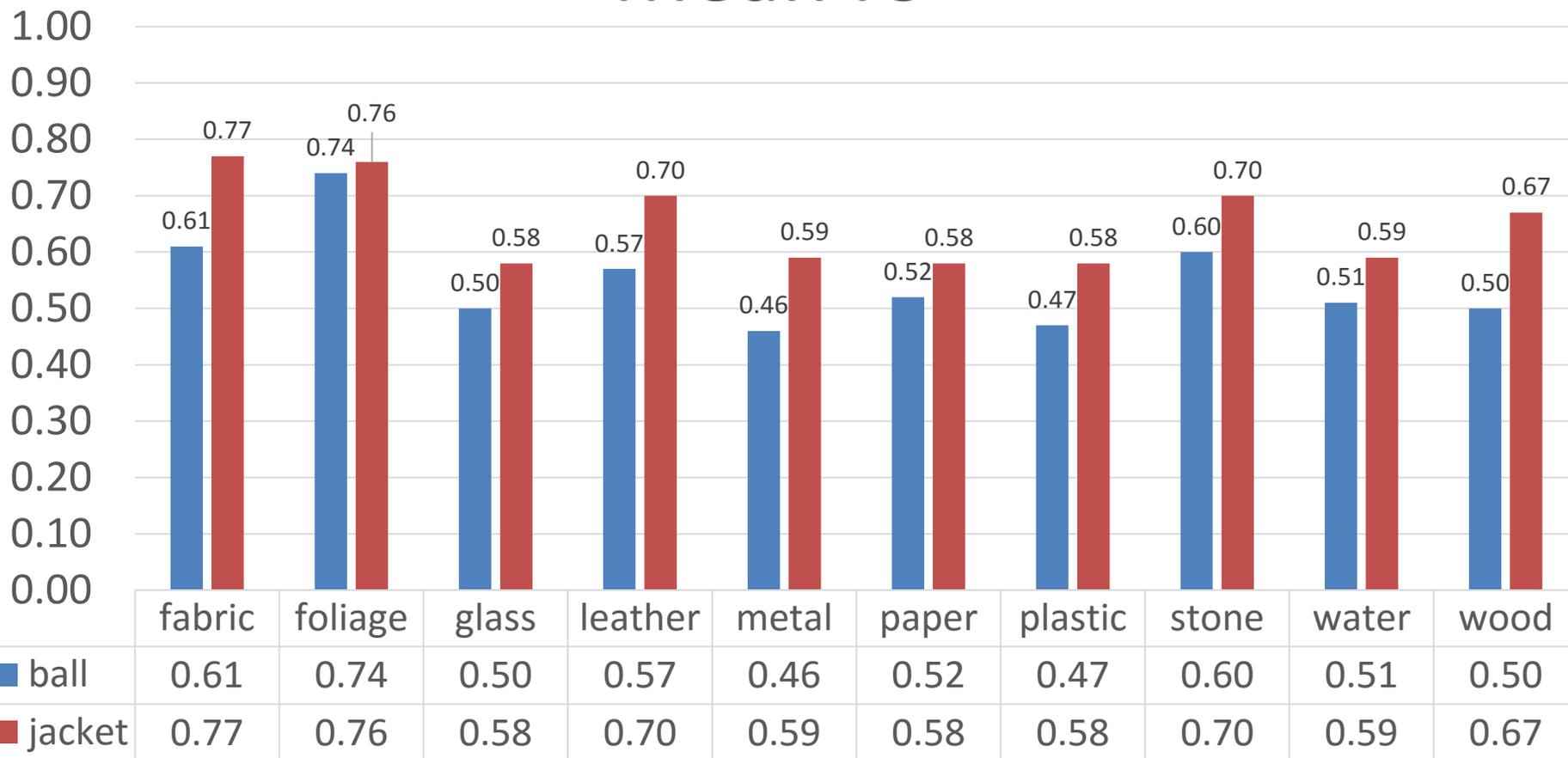
# 再領域分割結果の評価

## Pixel accuracy



# 再領域分割結果の評価

## Mean IU

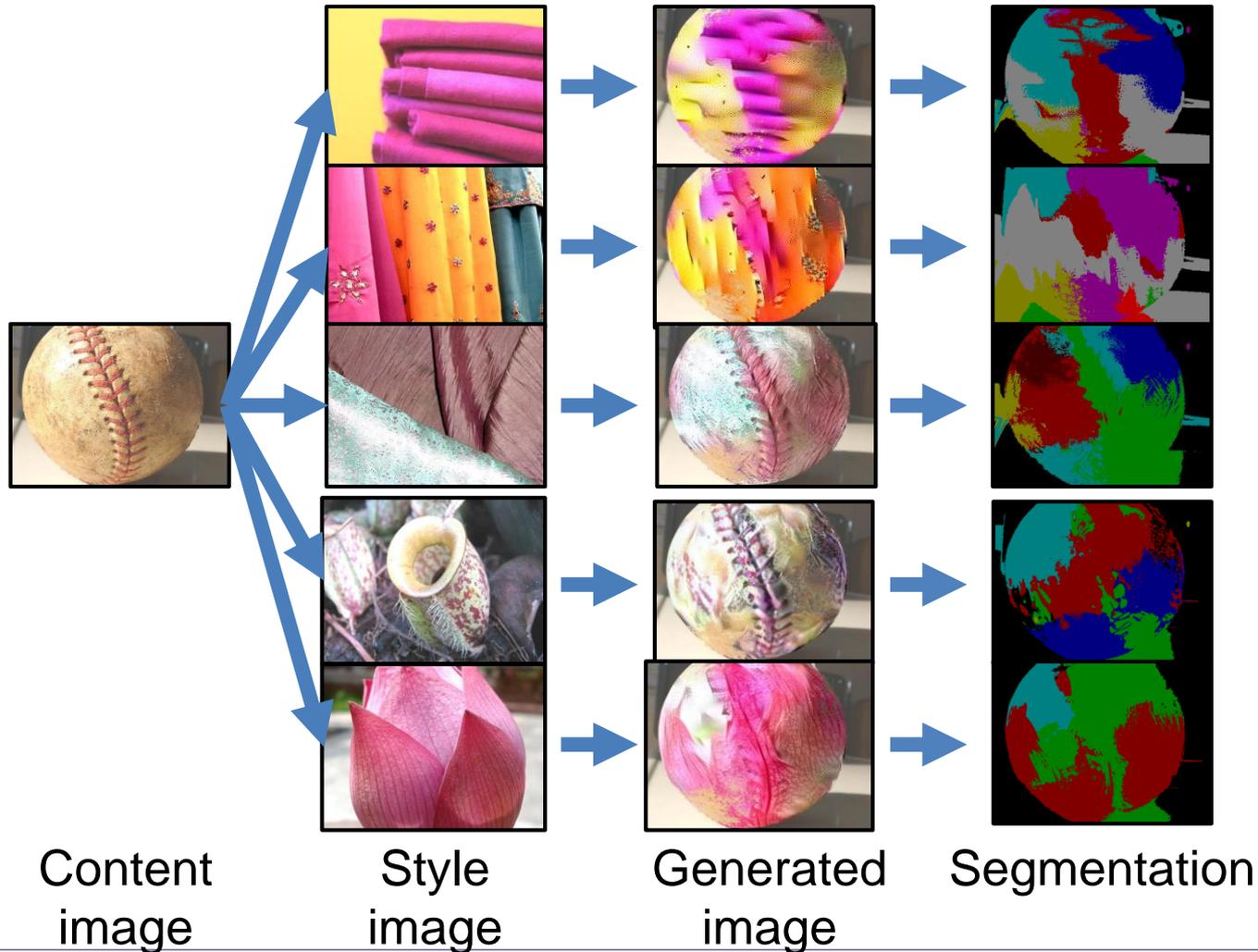


# 精度の高い素材について

foliage, fabric, stone素材による変換結果は精度が高かった。

- 不規則かつ多様な構造を持つ
  - 質感変換が成功しやすい
  
- 画像セット内の画像の構造に共通性が多い
  - 学習が成功しやすい

# 精度の高い素材について

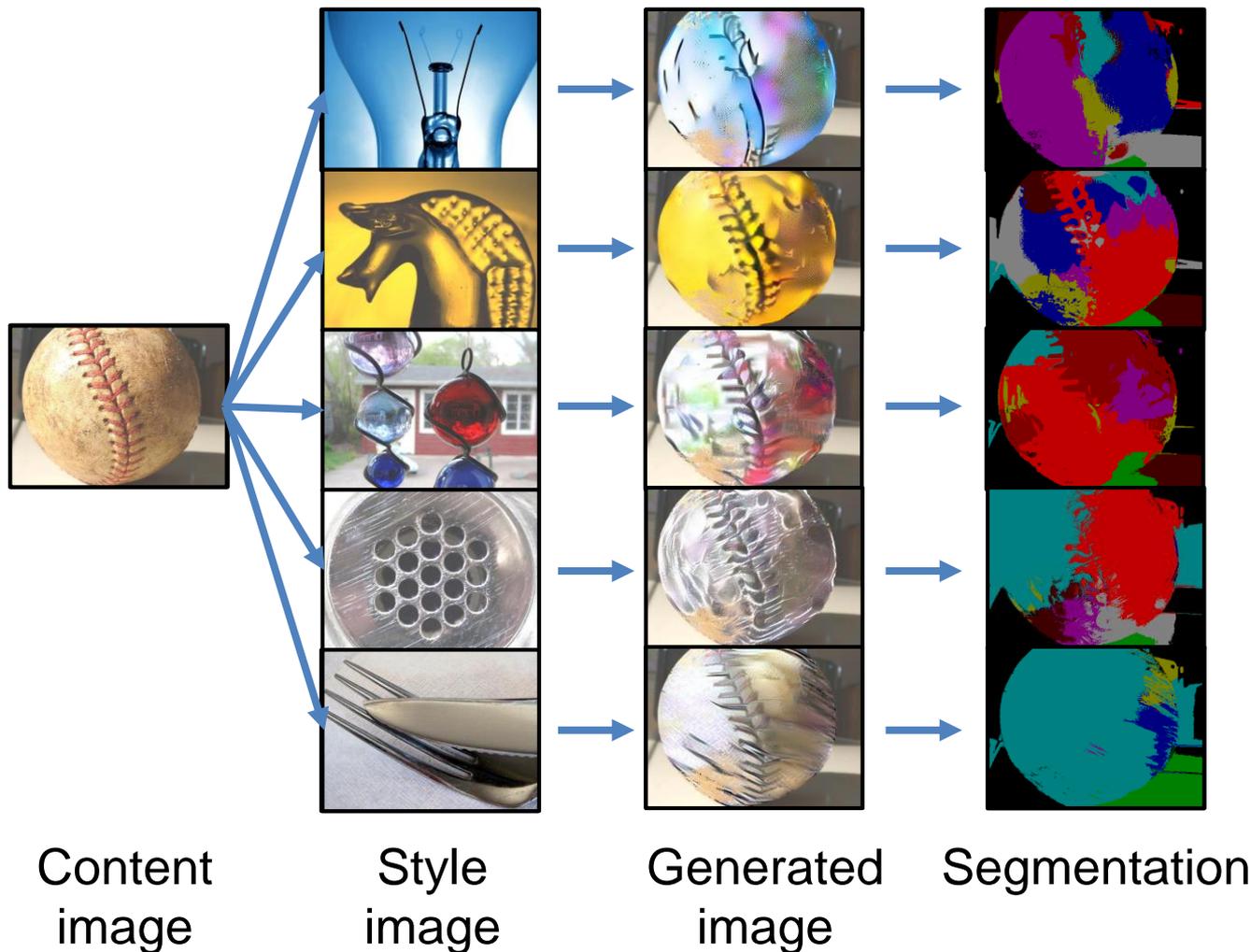


# 精度の低い素材について

glass, metal, plastic素材による変換結果は、精度が低かった。

- 規則的な構造が多い
  - 質感変換が成功しにくい
  
- 画像セット内の画像の構造に共通性がない
  - 領域分割の学習が難しい
  - 外部環境により変化するため、光沢は難しい？

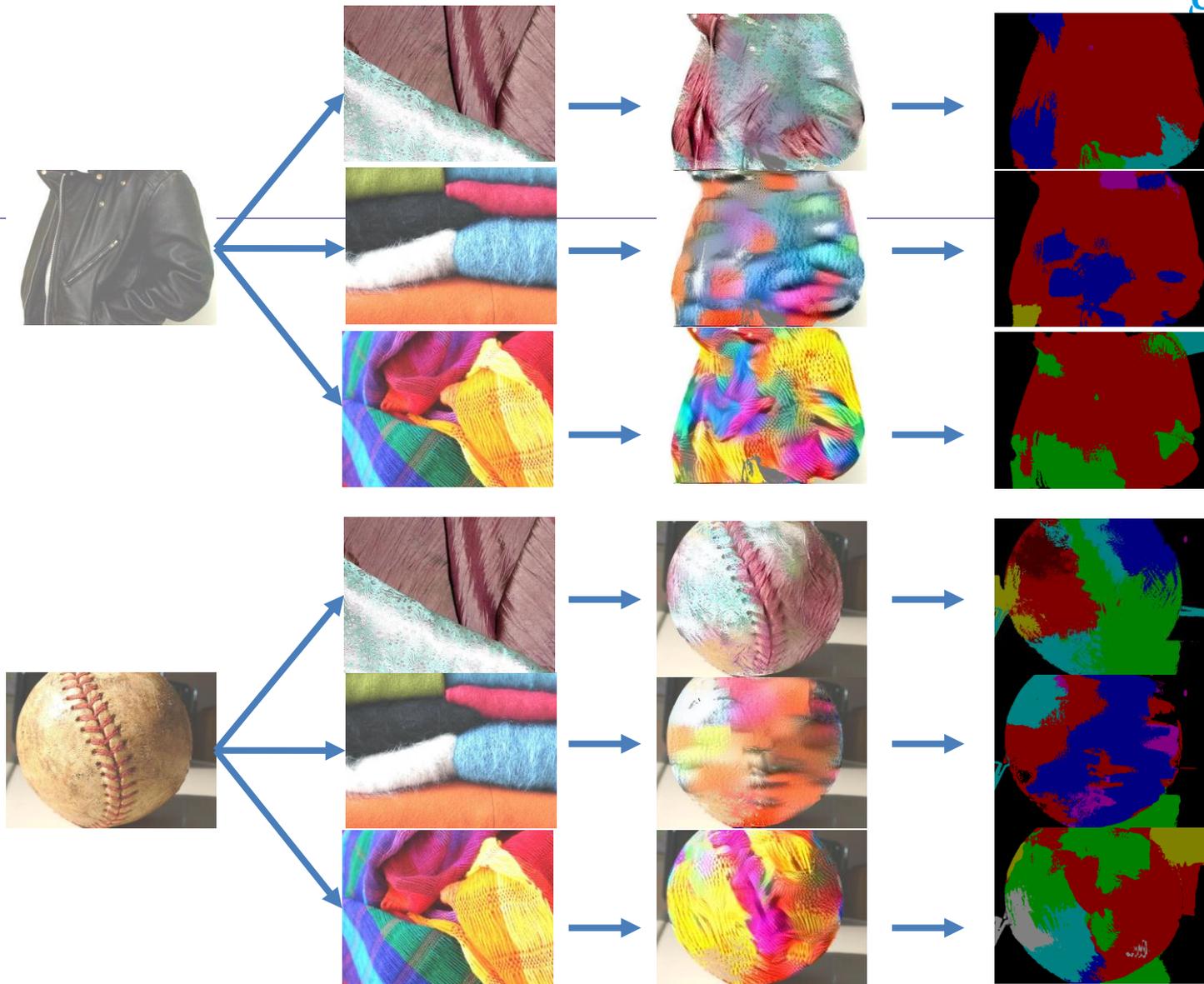
# 精度の低い素材について



# コンテンツによる精度の違い

## Fabricによる質感変換

- Jacketの領域分割の精度がballより高い  
 – 0.77 : 0.61 (Mean IU)
- Jacket画像とスタイル画像の共通する構造(しわ、しなり)が多いためか



Content image

Style image

Generated image

Segmentation

# まとめ

- DNNによる画像のスタイル変換 + 領域分割
  - スタイル変換アルゴリズムを画像内物体の質感の変換に応用
  - 違和感のない質感変換ができた。
  
- 素材ごとに質感変換の精度に違いが生じた。
  - 構造の多様性を持つ素材・コンテンツと共通構造を持つ素材はスタイル変換の精度を高める。

# 今後の課題

- スタイル変換の改良
  - 光沢を持つ質感への変換の改良
  - FMD以外の質感画像への応用
- ネットワークを一括化
  - 異なる手法の組み合わせを一つのネットワークにまとめる
  - 領域分割とスタイルの変換を同時に

# オノマトペ画像の学習

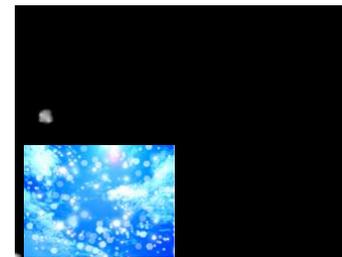
- ふわふわ、もふもふ、きらきら、ごつごつ、ぶつぶつ、ざらざら
- 6種類、200枚ずつ



ふわふわ



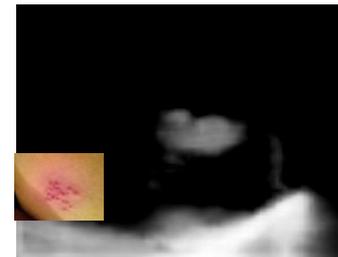
もふもふ



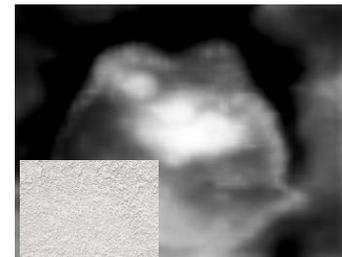
きらきら



ごつごつ



ぶつぶつ



ざらざら

# オノマトペによる質感変換

- 「もふもふ」の部位を「ごっごっ」に変換

