

学部卒業研究

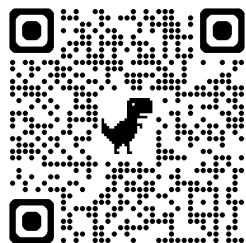
研究室ゼミ

中間発表



画像認識タスクに汎用な 熱赤外線画像着色モデルの検討

General-Purpose Infrared Image Coloring Model
for Various Recognition Tasks



[Any feedback](#)
is welcome!

大町・宮崎研究室
学部4年 谷内寛人

Hiroto TANIUCHI, 4th year undergraduate,
Laboratory for Image Information Communications ([IICLab](#)),
Tohoku University

目次

1. 導入 —赤外線画像の着色—
2. 関連研究 —宇川モデル—
3. 宇川モデルの発展性, 研究の動機
4. 研究方針, 提案手法
5. 実験計画
6. 結論
7. 話題

1.導入 ー赤外線画像の着色ー

“色彩”について研究したい！
...環境変化に強い頑健なセンサとして、熱赤外線カメラが注目されている

特徴

物体の熱を捉える

- ☑照明を必要としない=夜でも使える
- ☑悪天候でも使用可能



可視光画像

https://www.mi.t.u-tokyo.ac.jp/static/projects/mi_multispectral/det_result.png



熱赤外線(TIR)画像

応用例



自動運転^[1]

https://www.flir.com/globalassets/news/120x628_autonomous_vehicle.jpg



救助活動^[1]

<https://www.flir.jp/globalassets/defense/solution-and-landing-pages/ui/secure-arch-researcher.jpg/constrain-1130x0-2010510668.jpg>



防犯

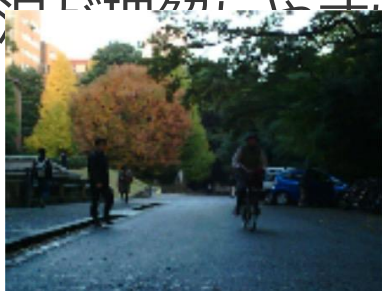
https://shop.flir.com/flir-cdn/shop/files/0000000776_k2HOT5k.png?v=1687514776

1.導入 ー赤外線画像の着色ー

可視光

✗ 環境の影響を強く受ける

✓ 状況が理解しやすい



https://www.mi.t.u-tokyo.ac.jp/static/projects/mi_multispectral/det_result.png

色情報テクスチャ

熱赤外線(TIR)

✓ 天候・照明の変化に頑健

✗ 視認性が低く状況理解が困難



https://www.mi.t.u-tokyo.ac.jp/static/projects/mi_multispectral/det_result.png

熱の画像情報

赤外線画像から生成した擬似可視光画像

✓ 状況が理解しやすい

✓ 天候・照明の変化に頑健

2. 関連研究 一宇川モデル

- ・ 大町・宮崎研究室2023年度卒業生の宇川さんが提案したモデル
- ・ TICC-GAN^[3](敵対的生成ネットワークの一つ)をベースラインに使用
- ・ セグメンテーションモジュールからの特徴マップを参照することで**物体の意味を適切に反映した着色画像を生成**



～2023/11/25
駅伝のおもひで～

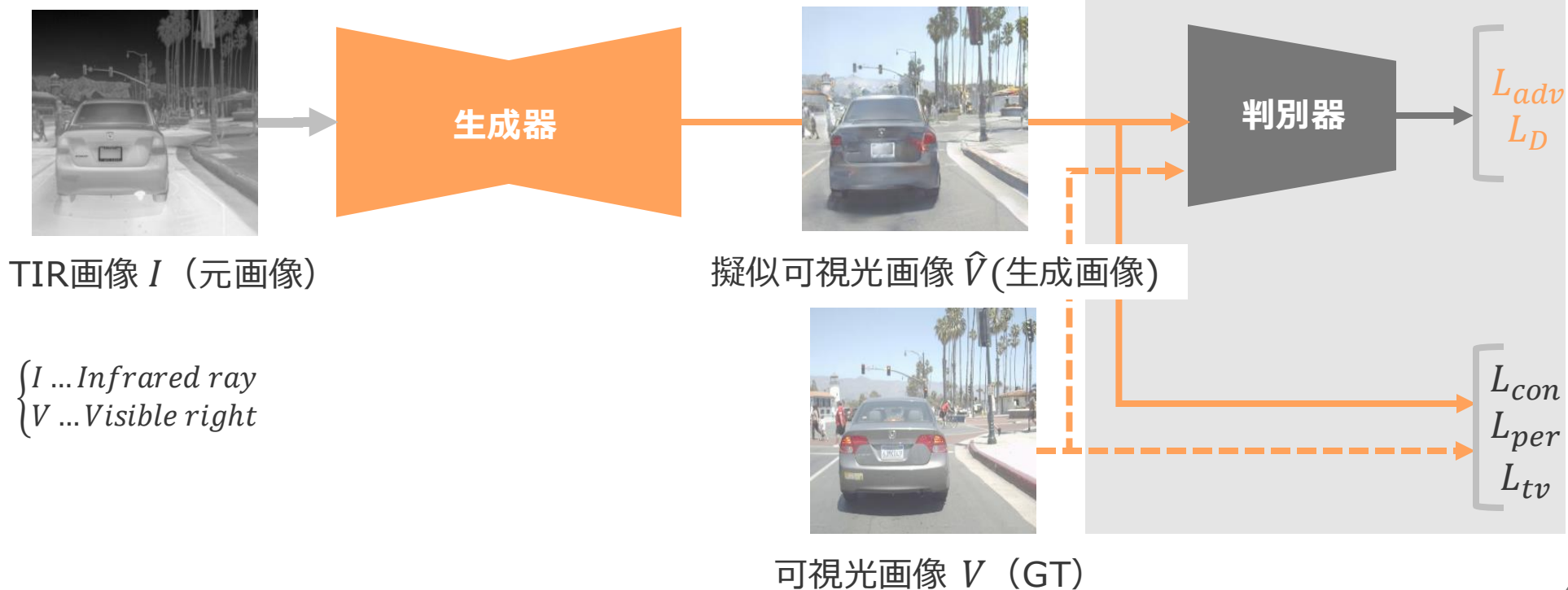
2. 関連研究 一宇川モデル

GANを利用した従来の着色モデル

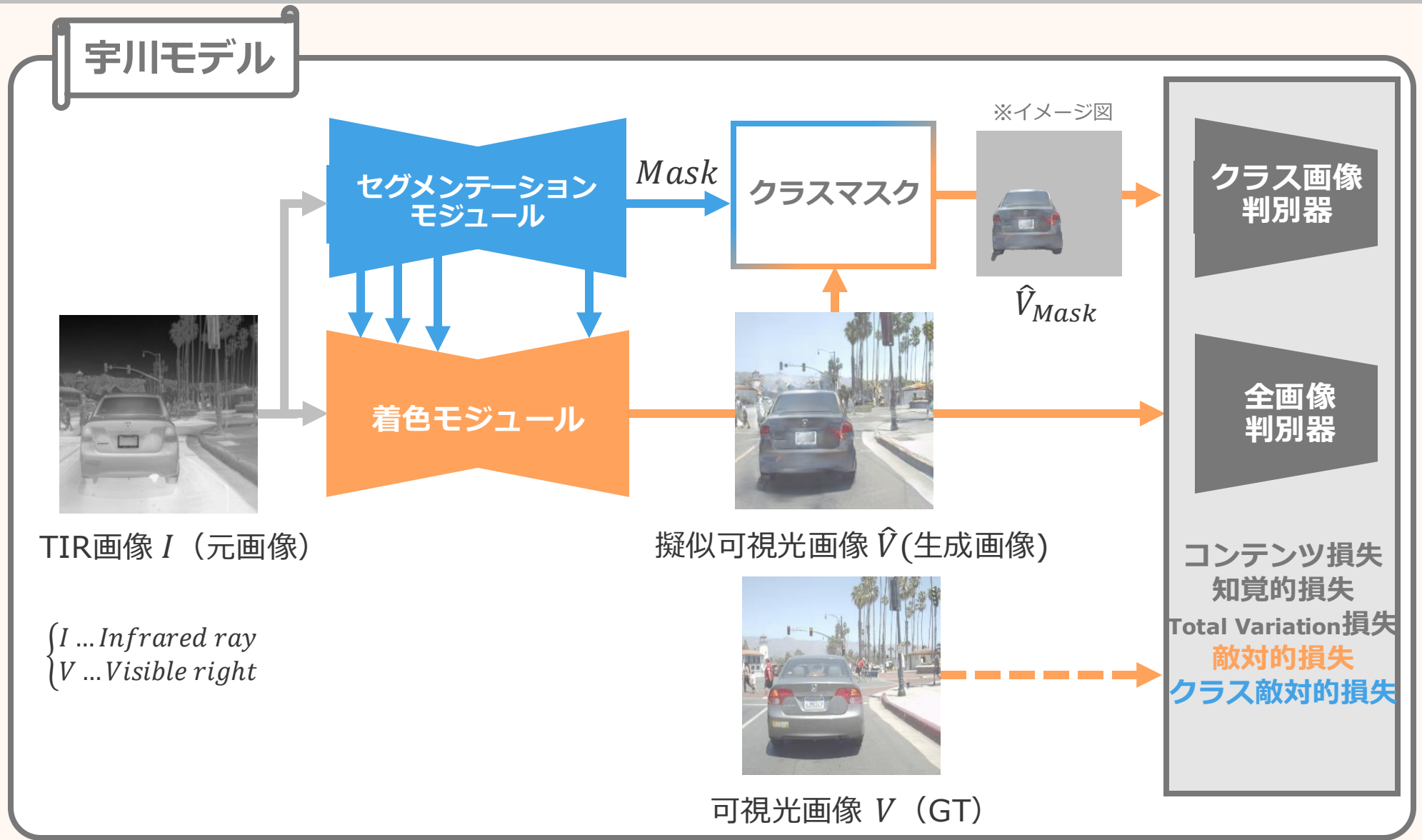


2. 関連研究 一宇川モデル

GANを利用した従来の着色モデル ~損失計算~



2. 関連研究 一宇川モデル

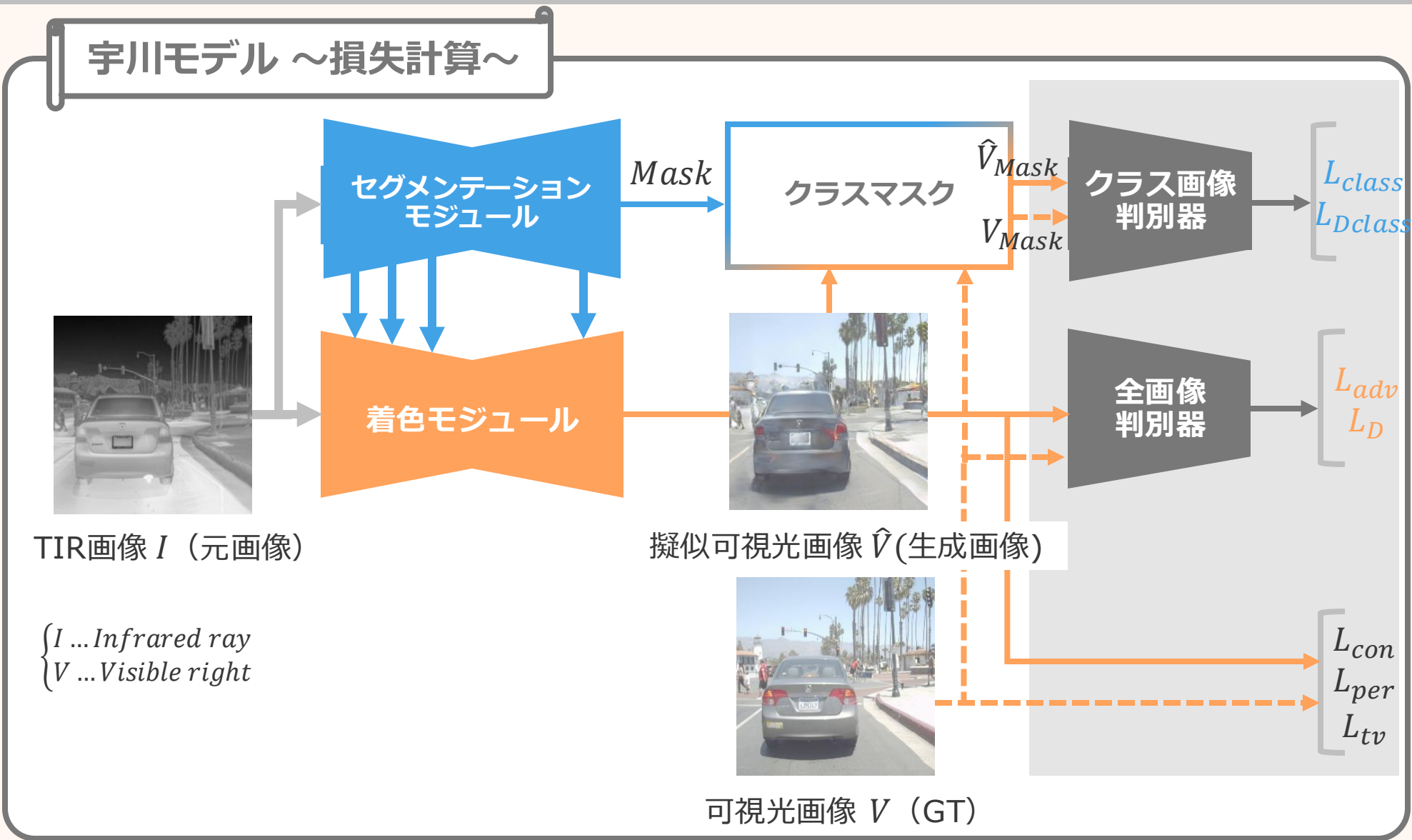


2. 関連研究 一宇川モデル

GANを利用した従来の着色モデル



2. 関連研究 一宇川モデル



2. 関連研究 一宇川モデル

従来の着色モデル ~損失計算~



3. 宇川モデルの発展性, 研究の動機

着眼点

- ・ 実験の一環として、出力した擬似可視光画像 \hat{V} を**可視光画像で学習したセグメンテーションモデル $Model^S$** に入力
- ・ 着色の精度が不足しており、正確な分類はできなかった
(そもそも宇川モデルは画像品質の向上を目的としている)
- ・ もし**可視光画像で学習済**のモデルに入力してもうまく動作するならば、公開されている訓練済みの大規模モデル(ResNet, BiT, CLIP...)を赤外線画像にも適用できる！
- ・ すなわち、“**人にとって**”だけでなく“**画像認識モデルにとって**”の自然さも考慮しながら変換を行いたい。

動機

宇川モデルで着色した画像 \hat{V} を、**可視光画像で訓練済みの画像認識モデル** (セグメンテーションや物体検出) の入力に利用したい

4. 研究方針, 提案手法

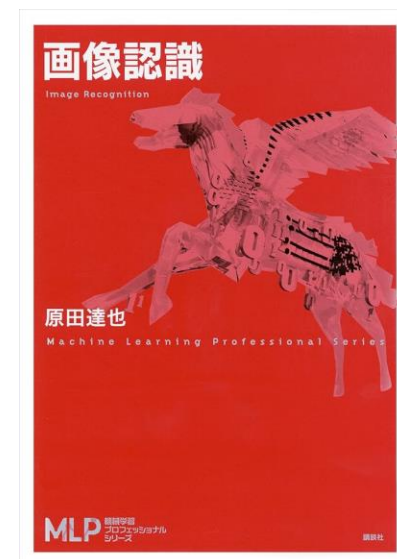
データセットは, 可視光画像・熱赤外線画像・アノテーション情報が同一シーン上で揃っていることが望ましい

MFNet データセット

- ・ 駐車場や道路の自然画像
- ・ 夜間及び昼間
- ・ 画素数640x480
- ・ 総数1606枚
- ・ 東京大学の原田・長・黒瀬・棕田研究室が作成^[4]

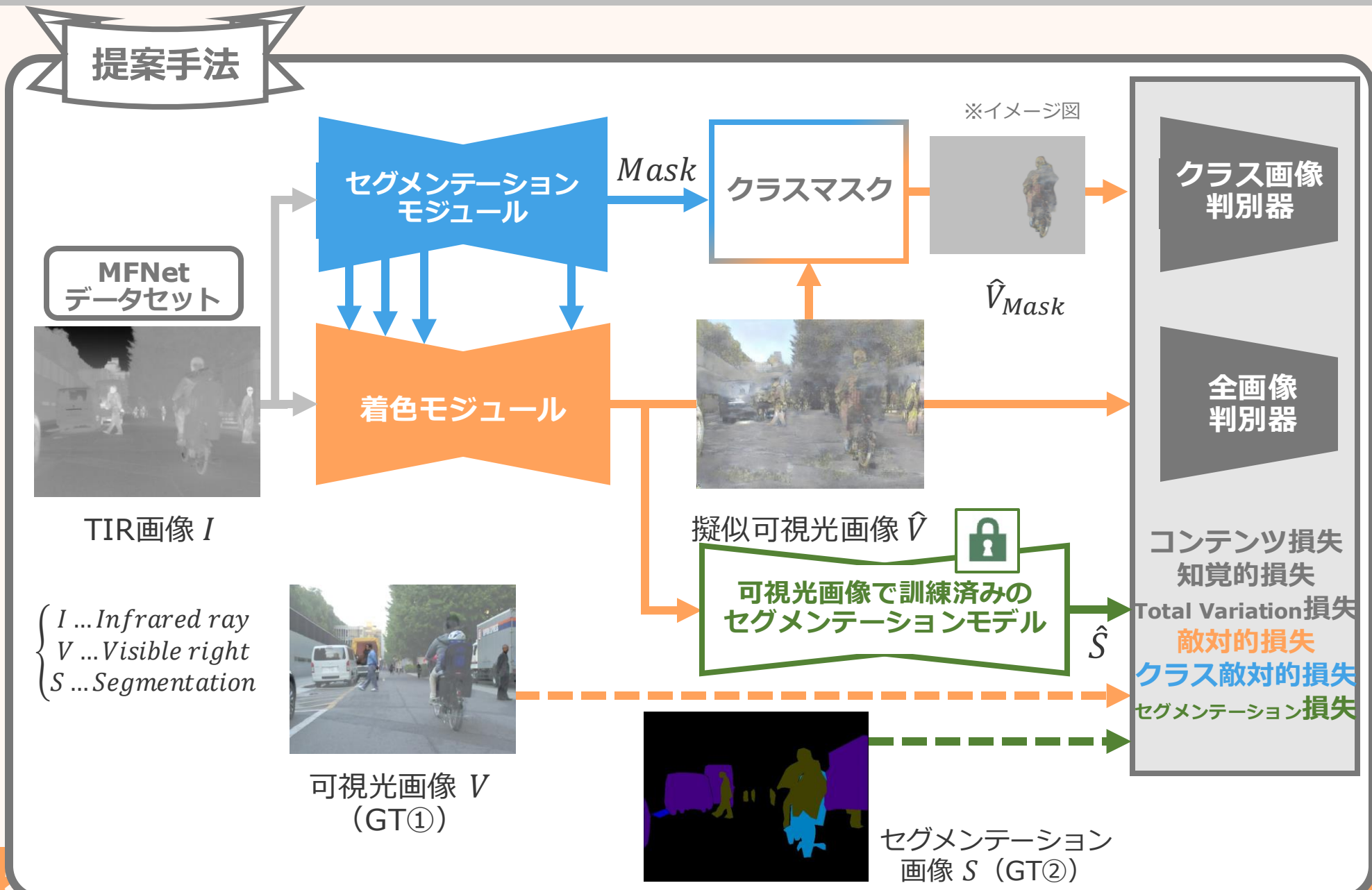


https://www.mil.t.u-tokyo.ac.jp/atic/proj/ctsg/mil_multispectral/predictionExample_s_good.png



https://m.media-amazon.com/images/I/91TCHGvJ2L._SY425_.jpg

4. 研究方針, 提案手法



4. 研究方針, 提案手法

GANを利用した従来の着色モデル

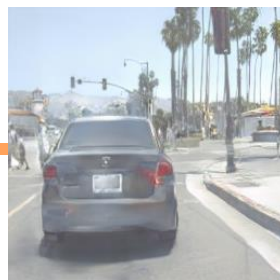


TIR画像 I

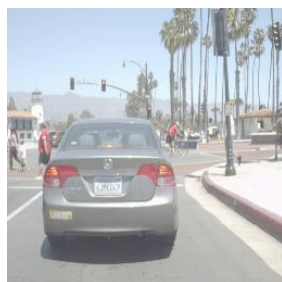
$\{I \dots \text{Infrared ray}$
 $\{V \dots \text{Visible light}$



生成器



擬似可視光画像 \hat{V}

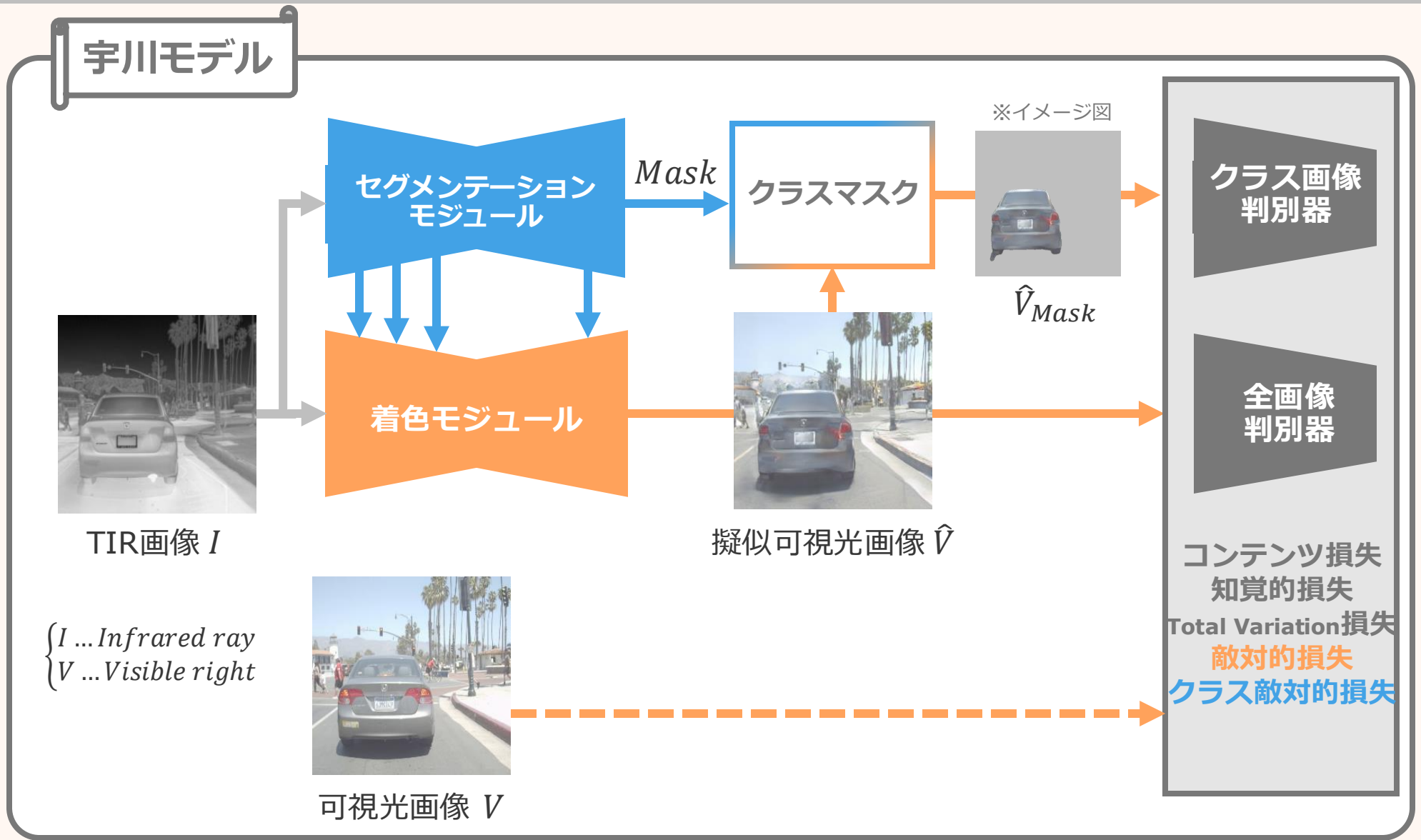


可視光画像 V

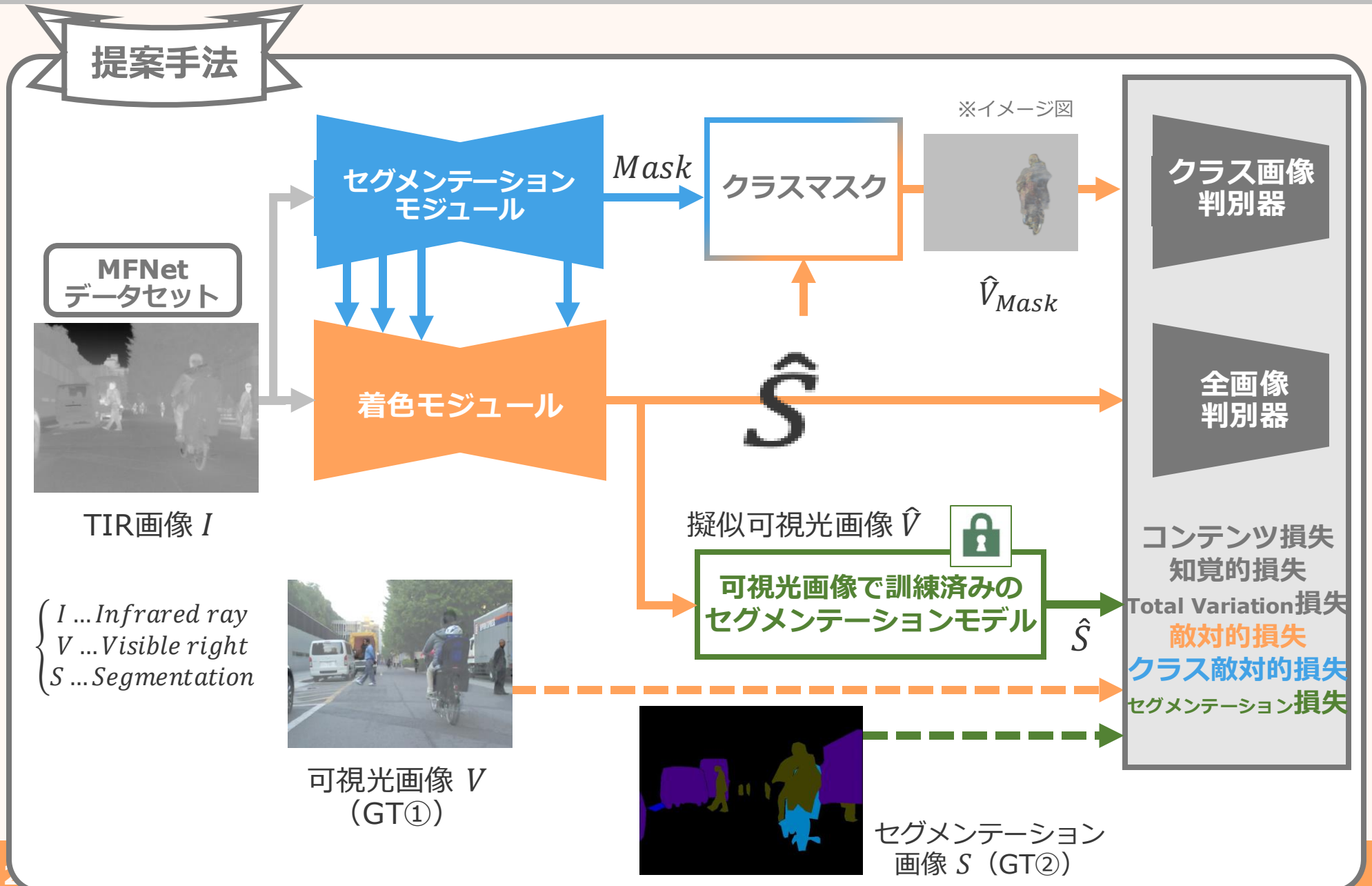
全画像
判別器

コンテンツ損失
知覚的損失
Total Variation損失
敵対的損失

4. 研究方針, 提案手法



4. 研究方針, 提案手法



5. 実験計画

- ☑ サーバに保存されている宇川モデルを動かす
- ☐ 宇川修論と同条件下で追実験を行い, 結果に近い値を自力で出す

12月 中間発表

- ☐ (宇川モデルを自力で実装する)
- ☐ MFNetデータセットの赤外線画像 I を訓練済の宇川モデル $Model^U$ に入力, 擬似可視光画像 \hat{V} を得る
- ☐ \hat{V} を**“一般可視光画像で訓練済の”**セグメンテーションモデル $Model^S$ に入力
本物の可視光画像 V を入力したときと分類精度を比較
= \hat{V} 入力時の出力 $s^{\hat{V}}$ は, V 入力時の出力 s^V より分類精度が低いことを確認

1月末

- ☐ 損失関数に「 \hat{V} 入力時の分類精度」を組み込んで $Model^{U'}$ を訓練
= “人間が見た際の自然さ”を維持した上で, **“一般可視光画像で学習済のモデルに入力してもうまく動作”**させるよう試みる

2月 最終発表

6. 結論

Motivation

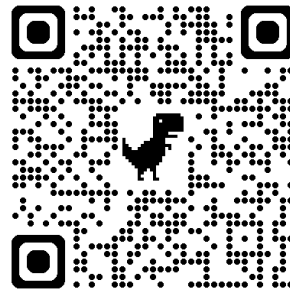
宇川モデルで着色した画像 \hat{V} を, 可視光画像で訓練済みの画像認識モデル (セグメンテーションや物体検出) の入力に利用したい

方針

- 判別器の損失関数に, “ \hat{V} をセグメンテーションした画像” \hat{S} とGT S との差 (=モデルにとっての不自然さ) を追加
- **同一シーン上の**可視光画像, 熱赤外線画像, セグメンテーション情報を持つデータセット, MFNetを導入
- いずれは物体検出タスクへの汎化も目指す

7. 話題

- 本研究へのアドバイスや、効果がありそうな手法
- モデルの性能に大きな影響を与えそうなパラメータ
- この研究テーマに新規性が認められるか
- エラー解決が辛い。泣。 • 実験条件の管理方法
- 既存のモデルを接続, 組み替える方法
- 画像認識モデルが何に注視しているかを検証する方法(Grad Cam?)
- GitHub上で公開されているリポジトリの探検の仕方
- GitHubによるバージョン管理の仕方 ...etc



[Any feedback](#)
is welcome!

ご清聴ありがとうございました



[Any feedback](#)
is welcome!

終
制作・著作



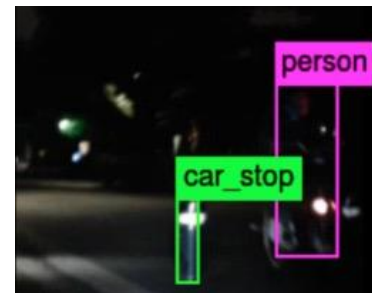
参考文献

- [1] **FLIR Systems** homepage <https://www.flir.jp/>
- [2] **Satoshi Ugawa**. A Study on Thermal Infrared Image Colorization Based on Semantic Information, 2024.
- [3] **X. Kuang et al.** 'Thermal infrared colorization via conditional generative adversarial network', Infrared Physics & Technology, vol. 107, p. 103338, Jun. 2020, doi: 10.1016/j.infrared.2020.103338.
- [4] **Qishen Ha, Kohei Watanabe, Takumi Karasawa, Yoshitaka Ushiku, Tatsuya Harada**. MFNet: Towards Real-Time Semantic Segmentation for Autonomous Vehicles with Multi-Spectral Scenes. The 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2017), 2017.
https://www.mi.t.u-tokyo.ac.jp/static/projects/mil_multispectral/

修士での展望, 研究アイデア

学部研究を継続

- ・ セグメンテーション以外のタスクへの汎化
= MFNetに含まれている**物体検出**データの利用
- ・ 宇川モデルの軽量化



https://www.mi.t.u-tokyo.ac.jp/static/projects/mil_multispectral/det_result.png

色彩に関連のある他のテーマ

- ・ カラー画像における色が持つ情報量...モノクロに変換すると失われる意味情報の大きさを評価する
- ・ パッケージ等の表示が色弱者にとってどの程度見やすいかを定量的に評価する
- ・ 気持ちやテーマを単語で入力すると, それに従ったスライドの配色を提案するサービスの実装

発表にて得られたフィードバック

- ・ GANよりも拡散モデルを利用した方が効果的な可能性

補助資料 ～宇川モデル 生成器の損失関数～

$$L_{all} = L_{con} + \lambda_{adv}L_{adv} + L_{per} + L_{tv} + \frac{\lambda_{class}L_{class}}{\text{損失の重み}}$$

- 既存手法^[2]で用いられている4つ損失関数に加えて、クラス敵対的損失を追加したものを生成器の学習に用いる損失関数とする

- コンテンツ損失 L_{con} : 生成画像とGround-truthの各画素の差分の絶対値 (L1損失)
- 敵対的損失 L_{adv} : 生成画像が判別器に(Ground-Truthよりも相対的に)本物らしいと判断されると小さくなる損失
- 知覚的損失 L_{per} : 生成画像とGround-truthを学習済VGGに入力した際の特徴マップのL1損失
- Total Variation損失 L_{tv} : 画像のx,y方向それぞれの勾配を損失とするもの
- クラス敵対的損失 L_{class}** : マスクをもとに特定クラスの領域を切り抜いた画像の敵対的損失

