



シヤムネットワークを用いたインタラクティブ認識システム

—♡☆♪♡◇♪◇♡☆◇♡♪☆♡◇☆—

徐 宏坤

2021/01/22

https://github.com/birdtianyu/LedgerMagic/tree/master/Interactive_Recognition_System

データセット

- “5-way 10-shot Learning” タスク

- ✓ 5カテゴリ
- ✓ 各カテゴリ10枚(256×256)
- ✓ 文字の種類ごとに筆の太さの違いがある。

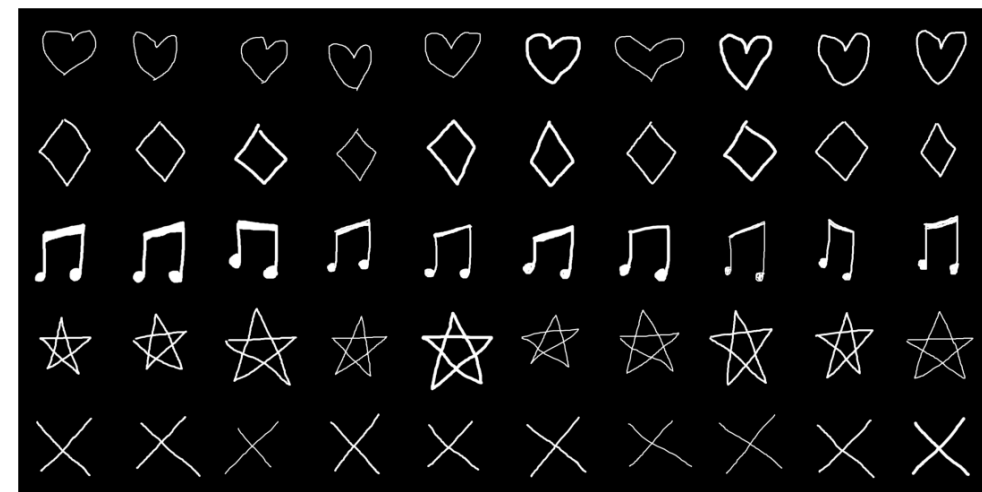
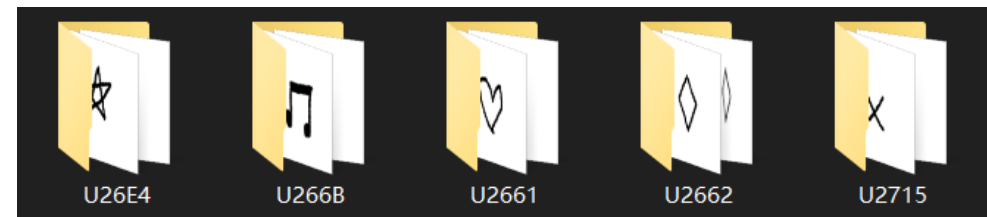


図 1: 全てのデータ

前処理

- ・シャムネットワーク訓練用データ

トレーニングサンプル:



ラベル: [0. 0. 1. 1. 0. 1. 0. 0. 1. 0.]

テストサンプル:



ラベル: [0. 0. 0. 1. 1. 0. 0. 0. 1. 0.]

図 2: シャムネットワークのトレーニングデータとテストデータの例

モデル

・シャムネットワーク

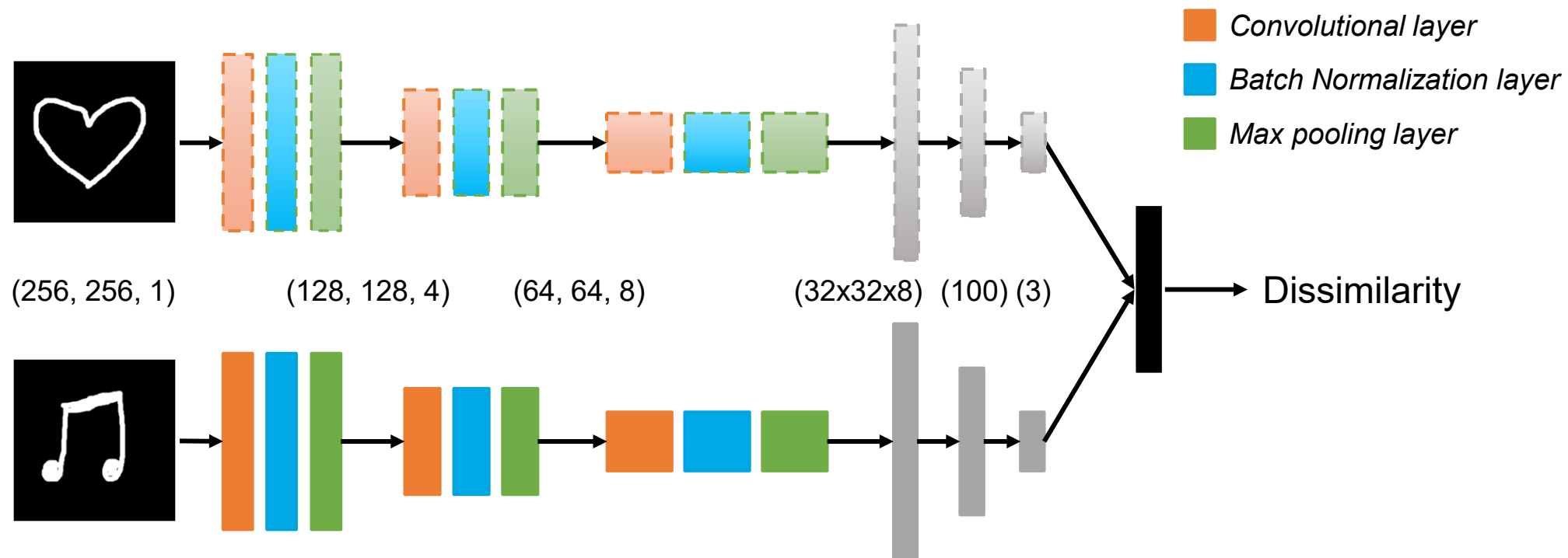


図 3: 認識システムで使用するシャムネットワークのアーキテクチャ

損失関数

- ユークリッド距離 (Euclidean distance)

$$D_W (X_1, X_2) = \|G_W (X_1) - G_W (X_2)\|_2$$

- コントラスト損失関数 (Contrastive Loss Function)

$$\mathcal{L}(W) = \sum_{i=1}^P L \left(W, (Y, X_1, X_2)^i \right)$$

$$L \left(W, (Y, X_1, X_2)^i \right) = Y L_S (D_W^i) + (1 - Y) L_D (D_W^i)$$

- 例:

$$L (W, Y, X_1, X_2) = Y (D_W)^2 + (1 - Y) \{ \max (0, m - D_W) \}^2$$

$$m = 2$$

結果

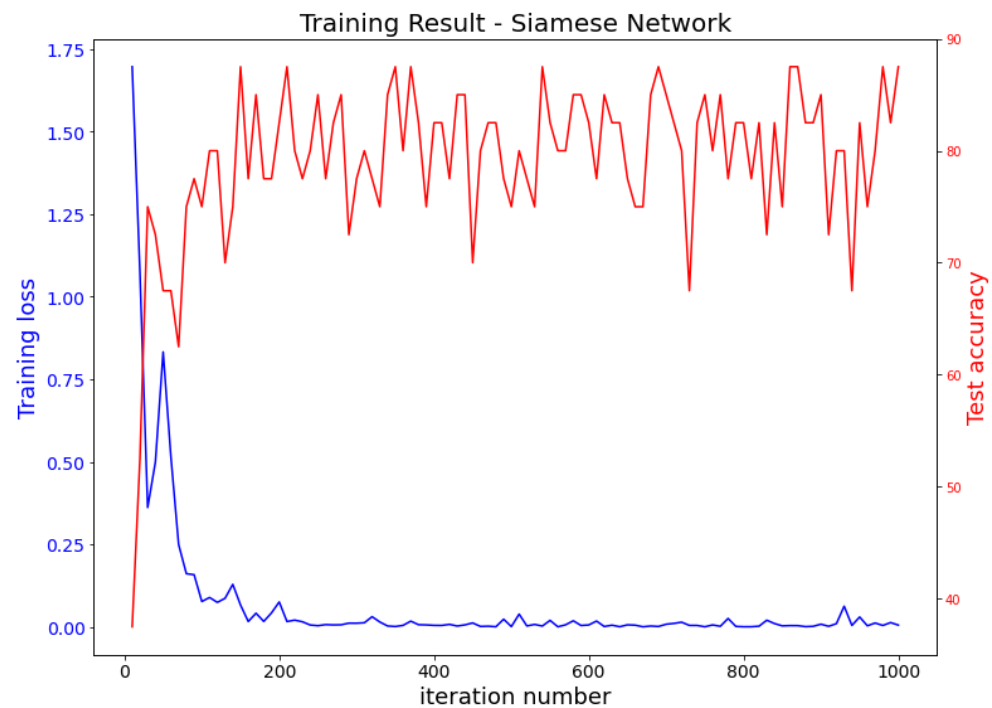


図 4: 損失値とテスト精度

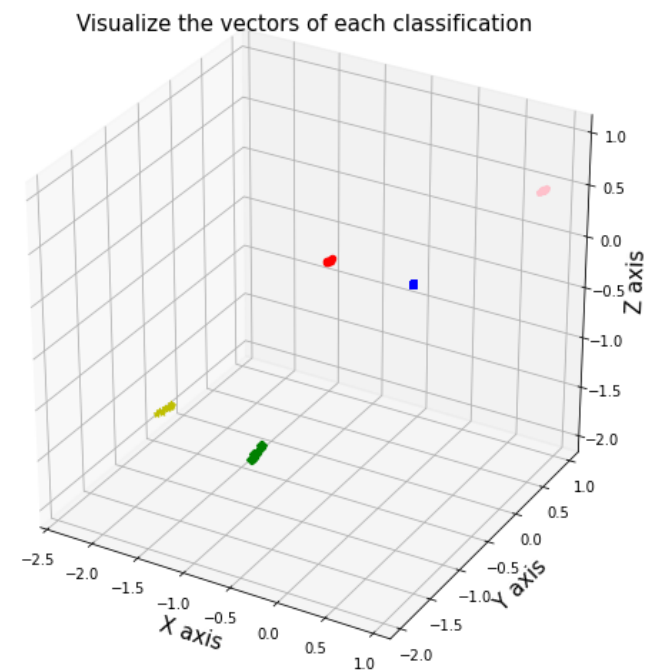


図 5: すべての学習データに対して出力されたベクトル値の3D表現

結果

- 非類似度 (Dissimilarity)

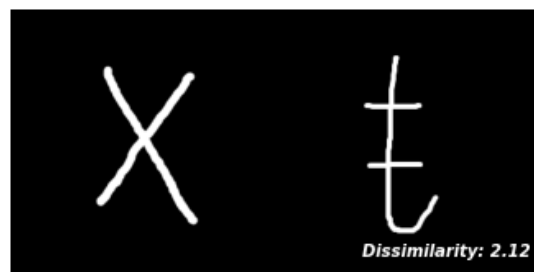
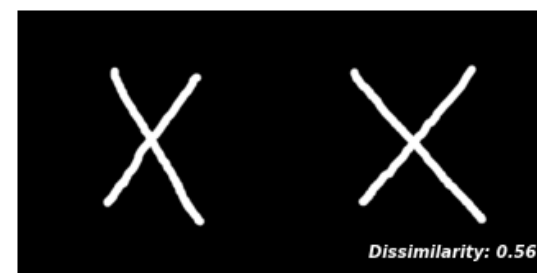
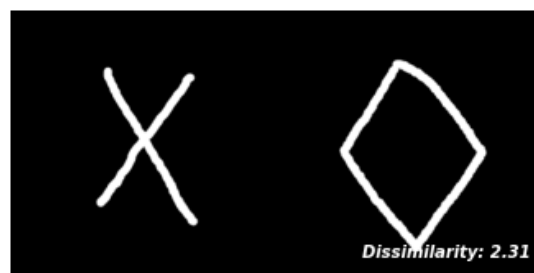


図 6: テスト画像の例.

インタラクティブ認識システム

✓ 各クラスを表すために、各カテゴリの出力の平均ベクトルを使用する。

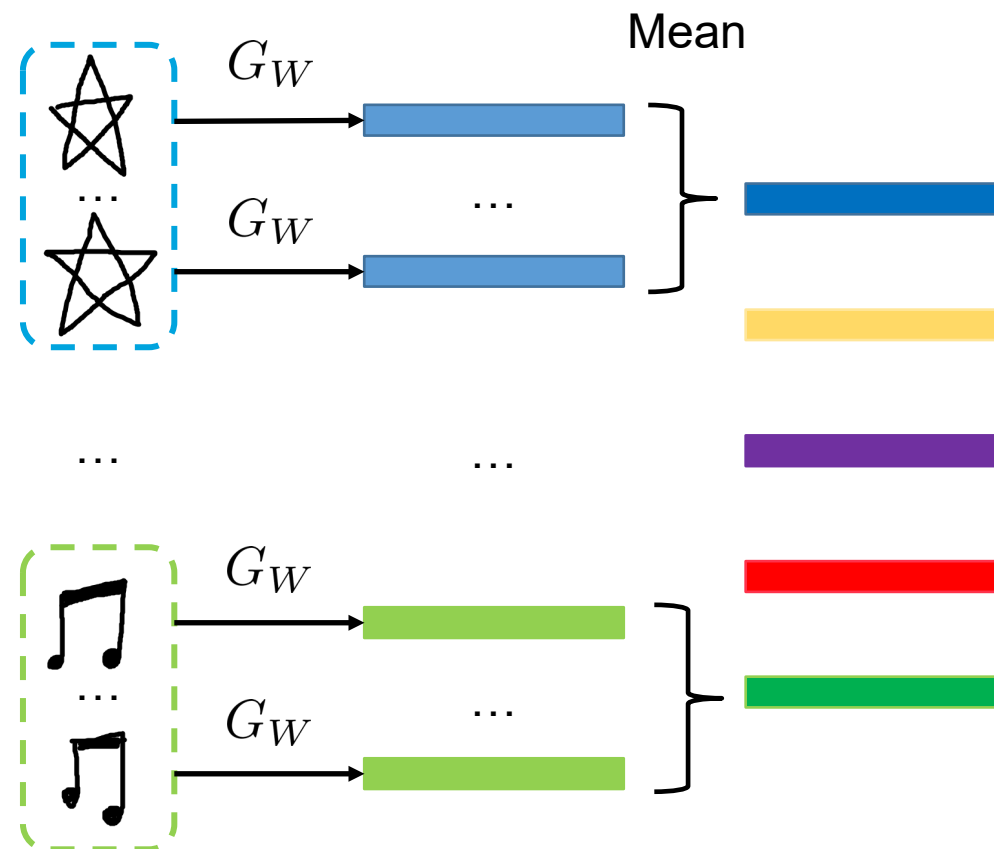


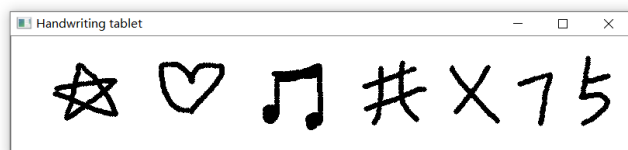
図 7: 平均ベクトル化.

インタラクティブ認識システム



インタラクティブ認識システム

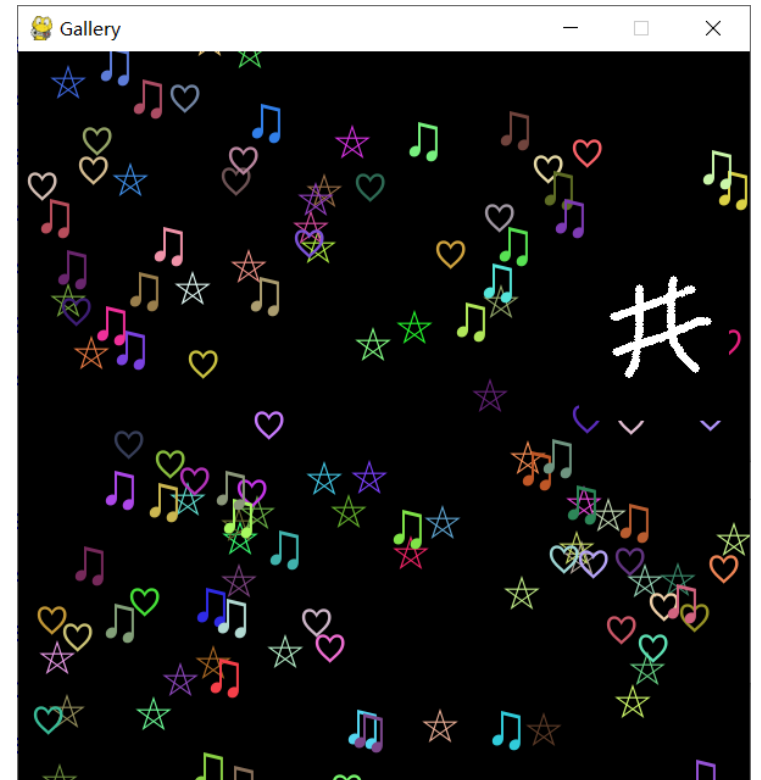
- ✓ 認識結果で効果を変えることができる
- ✓ 未知のカテゴリの画像もそれに対応する効果がある。



一緒に魔法使いになりましょう!



```
=====
part0, Siamese Prediction:3, Dissimilarity:0.332258403301239
CNNs Prediction:3, Probability:0.9093626141548157
part1, Siamese Prediction:0, Dissimilarity:0.30421799421310425
CNNs Prediction:0, Probability:0.929562509059906
part2, Siamese Prediction:2, Dissimilarity:0.8885497450828552
CNNs Prediction:2, Probability:0.757458508014679
part3, Siamese Prediction:2, Dissimilarity:1.3135663270950317
CNNs Prediction:2, Probability:0.794722855091095
part4, Siamese Prediction:4, Dissimilarity:0.19852982461452484
CNNs Prediction:4, Probability:0.9778745770454407
part5, Siamese Prediction:0, Dissimilarity:0.4483475387096405
CNNs Prediction:2, Probability:0.4441388249397278
part6, Siamese Prediction:0, Dissimilarity:0.7602939009666443
CNNs Prediction:0, Probability:0.5048049092292786
prediction: ['☆', '♡', '♪', '?', '✕', 7, 4]
Command: ['☆', '♡', '♪'] x 74 times
```



ご清聴ありがとうございました

