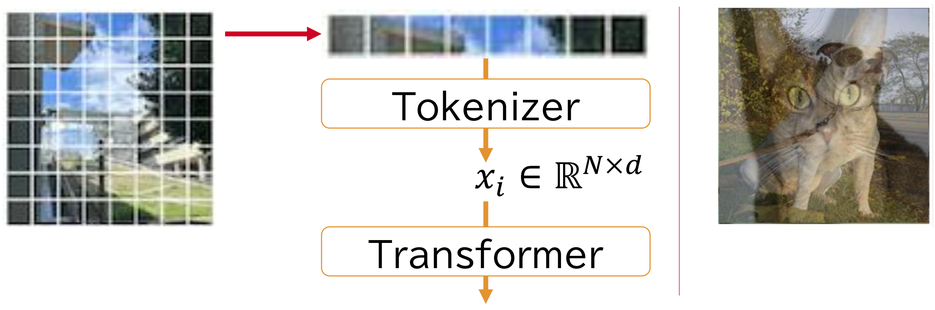
特徴部分空間の入れ替えによる データ拡張手法の提案

- ▶ 転移学習におけるデータ拡張手法の一種である Mixup は
  ２枚の画像を加重平均するため、一般的な画像では見られないような画像特徴が生じる
- ▶ そこで、加重平均ではなく、トークンの特定の次元の特徴量を交換することで得られる
  低次元表現を用いて学習・予測を行うDiSMiXを提案する

手法の概要

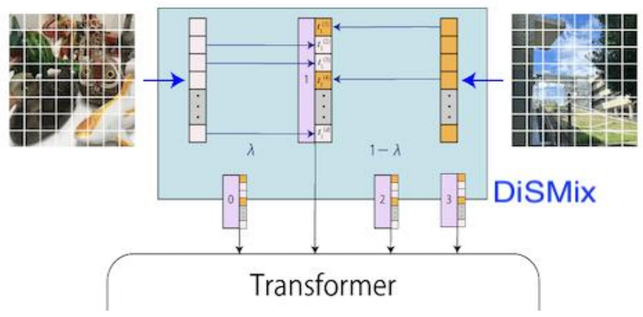
Overview of Vision Transformer (ViT)

- ■ ViT は画像からトークンを生成するTokenizerとTransformer からなる
- ■ Tokenizer：格子状に分割した画像の各領域の埋め込み表現を生成する
- ■ Transformer ： Attention を含むモジュールをレイヤー構造化したもの



DiSMix（our proposal）

- ■ Mixup などのデータ拡張では、学習データを増やすことで汎化性能の向上を目指す
- ■ DiSMix においても、一般的な手法と同様に、学習データから新たな学習例を生成する
  —— 全てのトークンにおいて一貫した特徴量の交換を行うことで、データを拡張する



$$\tilde{x} = Mx_1 + (I - M)x_2 \qquad M = \mathrm{diag}(b_1, b_2, \ldots, b_d) \qquad b \in \{0, 1\}$$

◆ 学習例（トークンとクラスラベルの組）を (x, y) とし、x_1 と x_2 はそれぞれ異なる画像を表す
◆ $\lambda \in \mathrm{Beta}(\alpha, \alpha)$ [ベータ分布] とし $\sum_i b_i \cong \lfloor \lambda d \rceil$ となるように決定する、本実験では $\alpha \in \{1.0, 0.5\}$
◆ y は正解ラベルにあたるクラスを表す one-hot ベクトルである　i.e. mixup データに対する Ground Truth は次式：$\tilde{y} = \lambda y_1 + (1 - \lambda) y_2$

結果のサマリー

Top-1 accuracy of VTAB-1k Natural

- ■ VTAB-1k Natural に属する sub-task の Top-1 精度
- ■ Flowers および Pets を除き, DiSMix が最も良い

method	Caltech	CIFAR	Flowers	Pets	Sun397	DTD	SVHN
ERM	87.3	77.3	**99.1**	87.3	40.7	70.7	89.5
mixup	88.2	78.7	98.9	**89.1**	41.3	72.3	90.5
Manifold	88.5	73.7	98.3	85.8	36.4	69.6	91.6
SMix	88.8	78.5	**99.1**	88.4	41.4	72.0	91.9
MSMix	88.4	73.1	98.3	86.0	36.4	70.7	91.7
DiSMix	**89.1**	**80.3**	99.0	88.9	**41.7**	**72.4**	**92.1**
MDiSMix	88.7	74.8	99.2	86.8	36.9	71.6	92.0

Average accuracy of VTAB-1k

- ■ VTAB-1k 全体の精度は下記の通り（詳細は割愛）

	VTAB-1k		
method	Natural	Specialized	Structured
ERM	78.3	85.8	56.0
mixup	79.3 (+1.0)	86.2 (+0.4)	56.2 (+0.2)
Manifold	77.1 (−1.2)	86.1 (+0.3)	58.6 (+2.6)
SMix	79.5 (+1.2)	86.1 (+0.3)	56.9 (+0.9)
MSMix	77.3 (−1.0)	86.1 (+0.3)	59.4 (+3.4)
DiSMix	**80.0** (+1.7)	86.1 (+0.3)	57.4 (+1.4)
MDiSMix	77.9 (−0.4)	**86.3** (+0.5)	**59.6** (+3.6)

Caltech/CIFAR: 汎用的なクラス　　Sun397 : シーン
Flowers/Pets: Fine-grained クラス　　DTD: テクスチャ　SVHN: 数字
Specialized: 鳥瞰図や医療画像　　Structured: 回転角、数、深度など

RiHSA 人間社会拡張研究部門
桐山
情報・人間工学領域 人間社会拡張研究部門
ソシオデジタルサービスシステム研究グループ
連絡先:M-rihsa-liaison-ml@aist.go.jp

25TH ANNIVERSARY
産総研
ともに挑む。つぎを創る。