AlignMixup: Improving Representations By Interpolating Aligned Features




Shashanka Venkataramanan

Ewa Kijak

Laurent Amsaleg

Yannis Avrithis

Mixup improves generalization

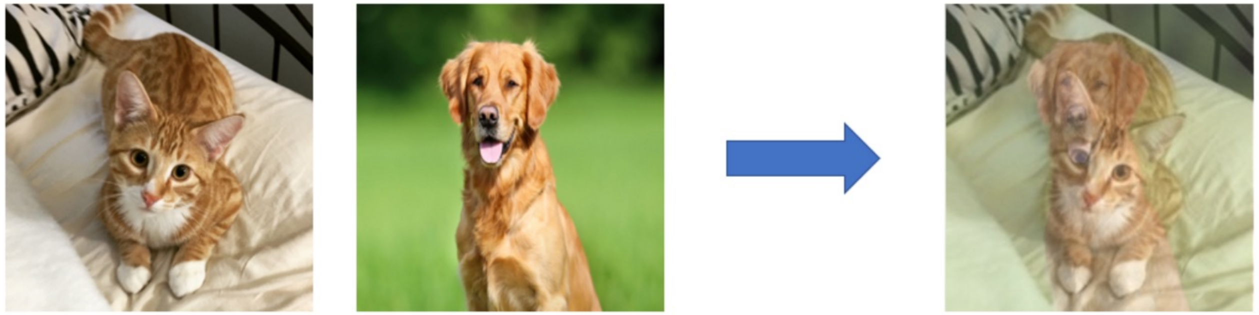
- Interpolates between pairs of examples (input/feature) and its target labels.

$$\lambda \sim \mathrm{Beta}(\alpha, \alpha)$$
$$\mathrm{mix}_{\lambda}(a, b\ ) = \lambda \mathrm{a} + (1 - \lambda) b'$$

[Zhang *et al.*, ICLR 2018]

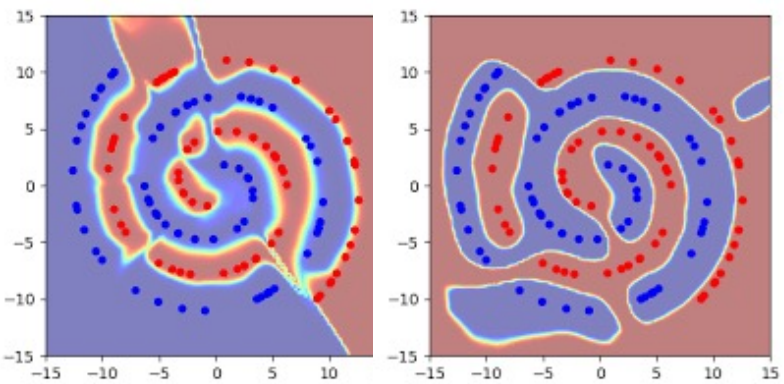
Mixup improves generalization

- Interpolates between pairs of examples (input/feature) and its target labels.
- Flattens class representations, reduces overconfident incorrect predictions, and smoothens decision boundaries.



[1.0, 0.0] cat dog   [0.0, 1.0] cat dog   [0.7, 0.3] cat dog

$$\lambda \sim \mathrm{Beta}(\alpha, \alpha)$$
$$\mathrm{mix}_{\lambda}(a, b) = \lambda a + (1 - \lambda) b'$$



without mixup   with mixup

[Zhang *et al.*, ICLR 2018; Verma *et al.*, ICML 2018]

Existing mixup methods

Co-Mixup

interpolates between the best combination of salient regions.

optimization is computationally expensive.

SaliencyMix

interpolates between an image patch computed using saliency with the target image.

images are unnatural and an overlay of one image onto another

[Kim *et al.*, ICLR 2021; Uddin *et al.*, ICLR 2021]

What is a good interpolation of images?

What is a good interpolation of images?

AlignMixup

AlignMixup: natural way of interpolation using deformation

- MOTIVATION - Deformation a natural way of interpolating images, one image may deform into another, in a continuous way.

- Interpolated points that smoothly traverse the underlying manifold, capture salient characteristics.

source target

[Berthelot *et al.,* ICLR 2019; Bengio *et al.,* 2013]

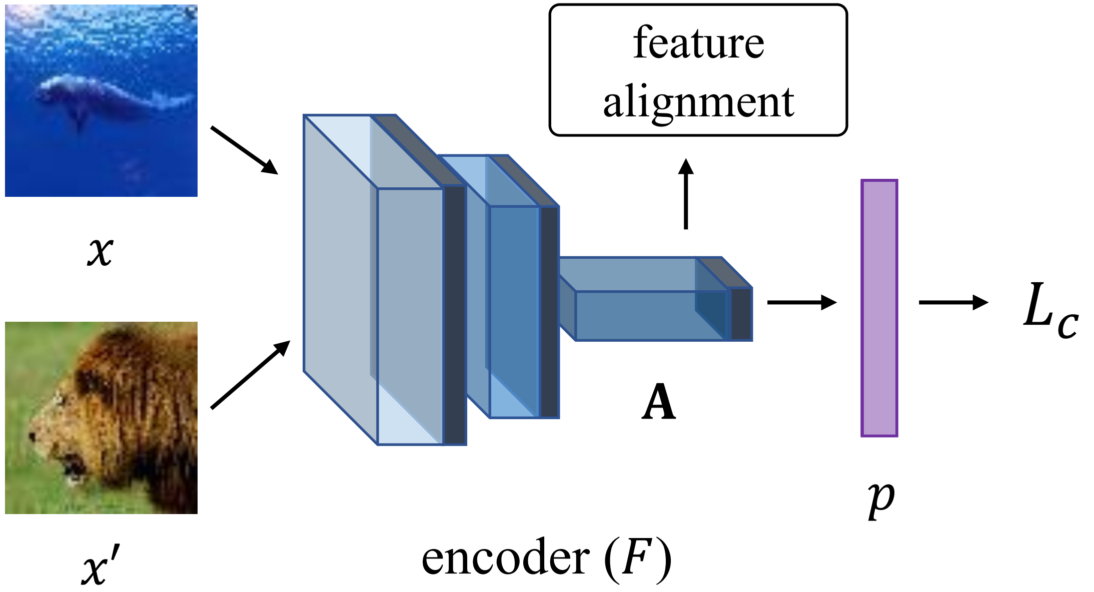
AlignMixup: interpolating aligned features

- Investigate geometric alignment for mixup, based on semantic correspondences in the feature space.

- Aligning features results in learning invariances [Choy et al., NIPS 2016].

- Deform objects across classes essentially populating the feature space between manifolds.

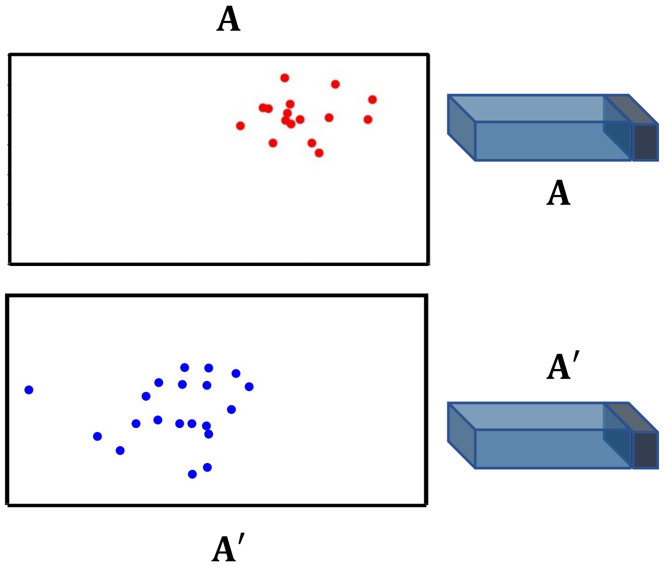
AlignMixup: interpolating aligned features



$$\mathbf{A} := F(x)$$

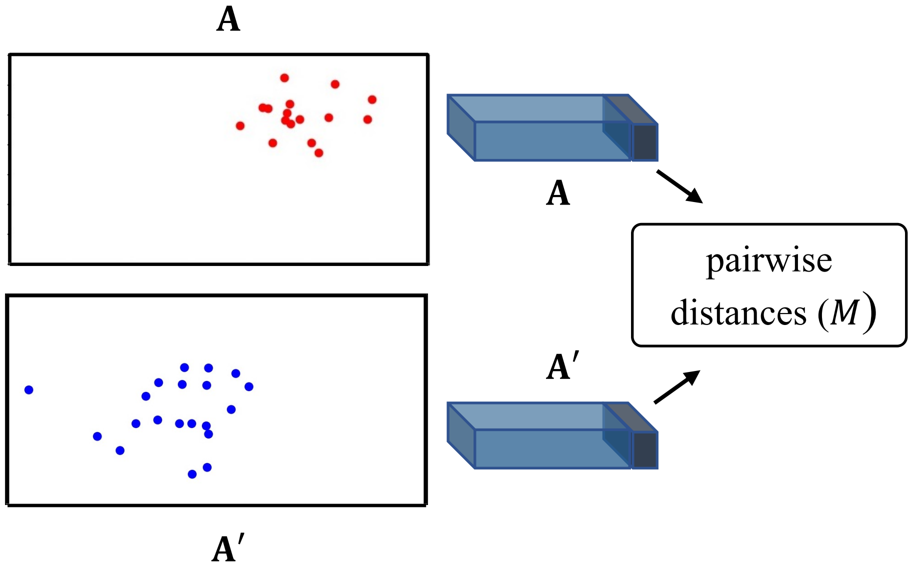
L_c is cross entropy loss

AlignMixup: aligning feature tensors



$$\mathbf{A} \coloneqq F(x);\ \mathbf{A}' \coloneqq F(x')$$

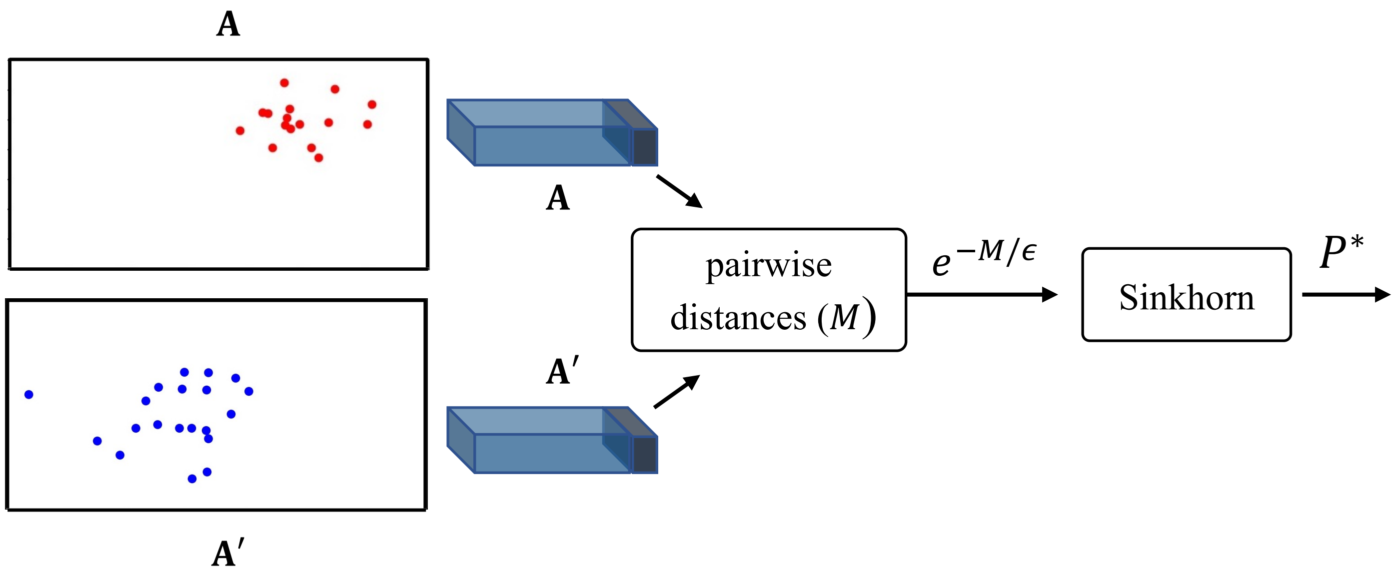
AlignMixup: aligning feature tensors



$$m_{ij} = \left|\left|a_i - a_j\right|\right|^2$$

m_{ij} is an element of M

AlignMixup: aligning feature tensors

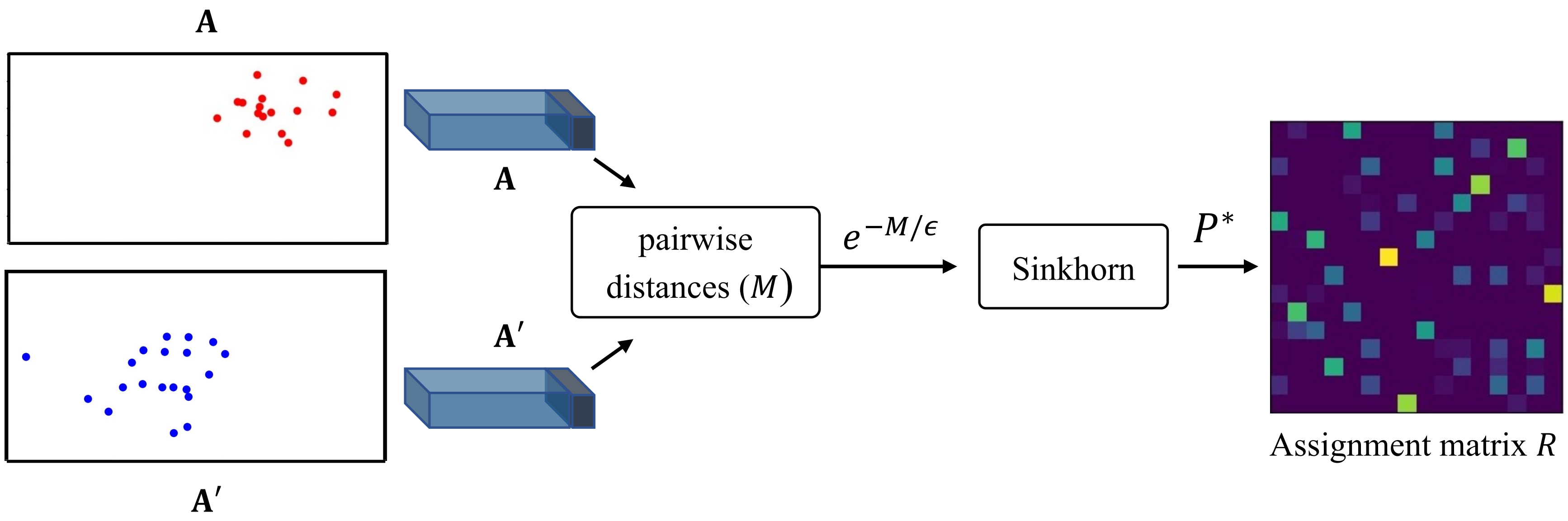


$$P^* = argmin_P \langle P, M\rangle - \epsilon H(P)$$

M is cost matrix

$H(P) = -\sum_{i,j} p_{ij} \log p_{ij}$ is entropy of P

AlignMixup: aligning feature tensors

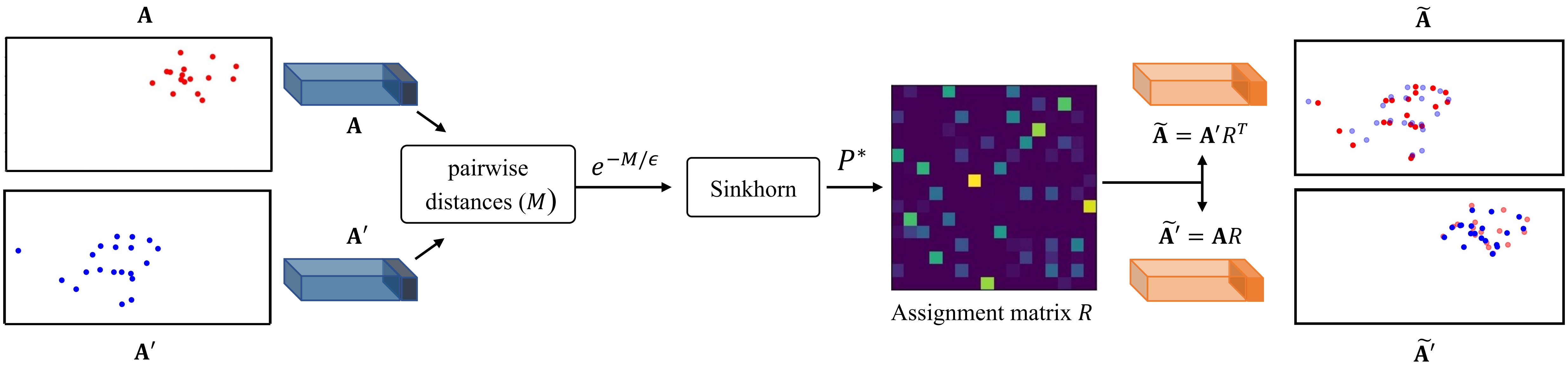


Assignment matrix R

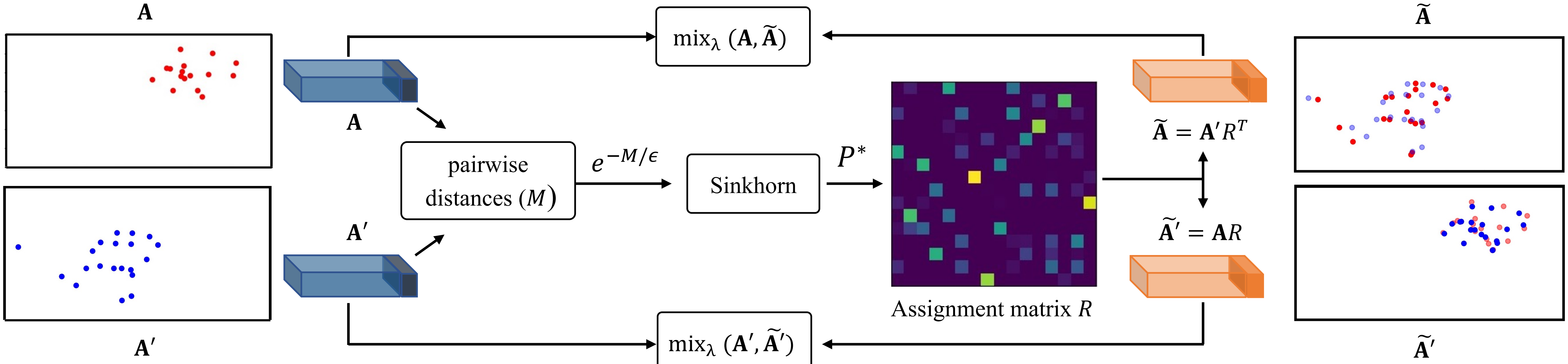
$$R = rP^*$$

$r = h \times w$ of feature A (A')

AlignMixup: aligning feature tensors



AlignMixup: interpolating aligned feature tensors



$$\mathrm{mix}_\lambda\left(\mathbf{A}, \widetilde{\mathbf{A}}\right) = \lambda\mathbf{A} + (1-\lambda)\,\widetilde{\mathbf{A}}$$

$$\mathrm{mix}_\lambda\left(\mathbf{A}', \widetilde{\mathbf{A}}'\right) = \lambda\mathbf{A}' + (1-\lambda)\,\widetilde{\mathbf{A}}'$$

AlignMixup: visualizing alignment

x x'

0.0 0.2 0.4 0.6 0.8 1.0 0.0 0.2 0.4 0.6 0.8 1.0 λ

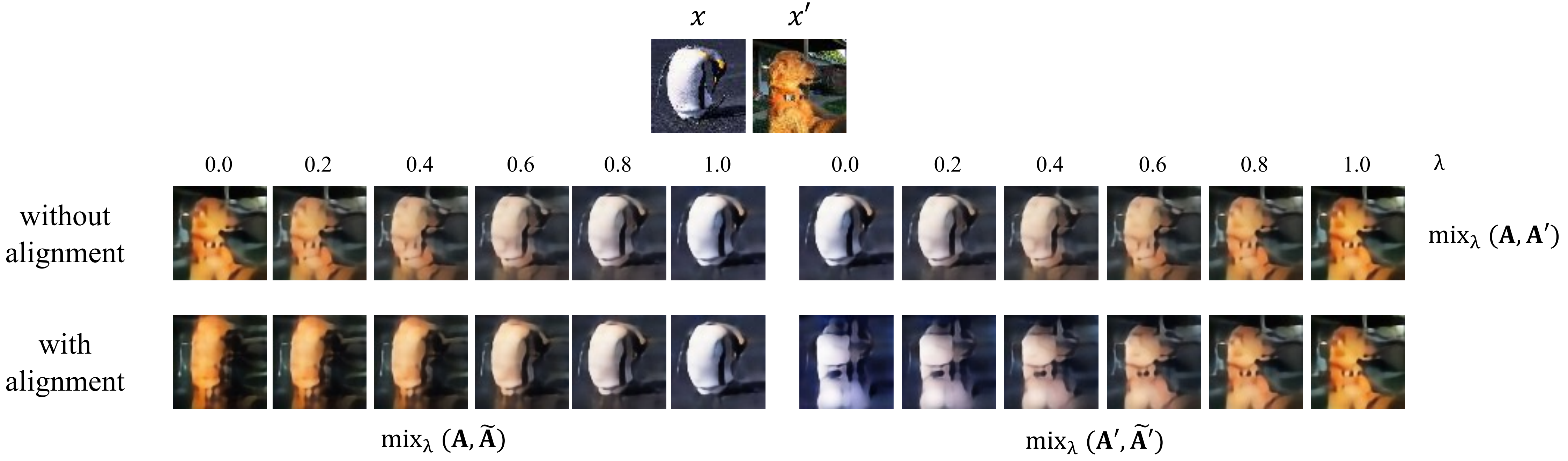
without alignment

$\mathrm{mix}_\lambda\ (\mathbf{A}, \mathbf{A}')$

$$\mathrm{mix}_\lambda\ (\mathbf{A}, \mathbf{A}') = \lambda \mathbf{A} + (\mathbf{1} - \lambda)\mathbf{A}'$$

[images only for visualization; not used for training]

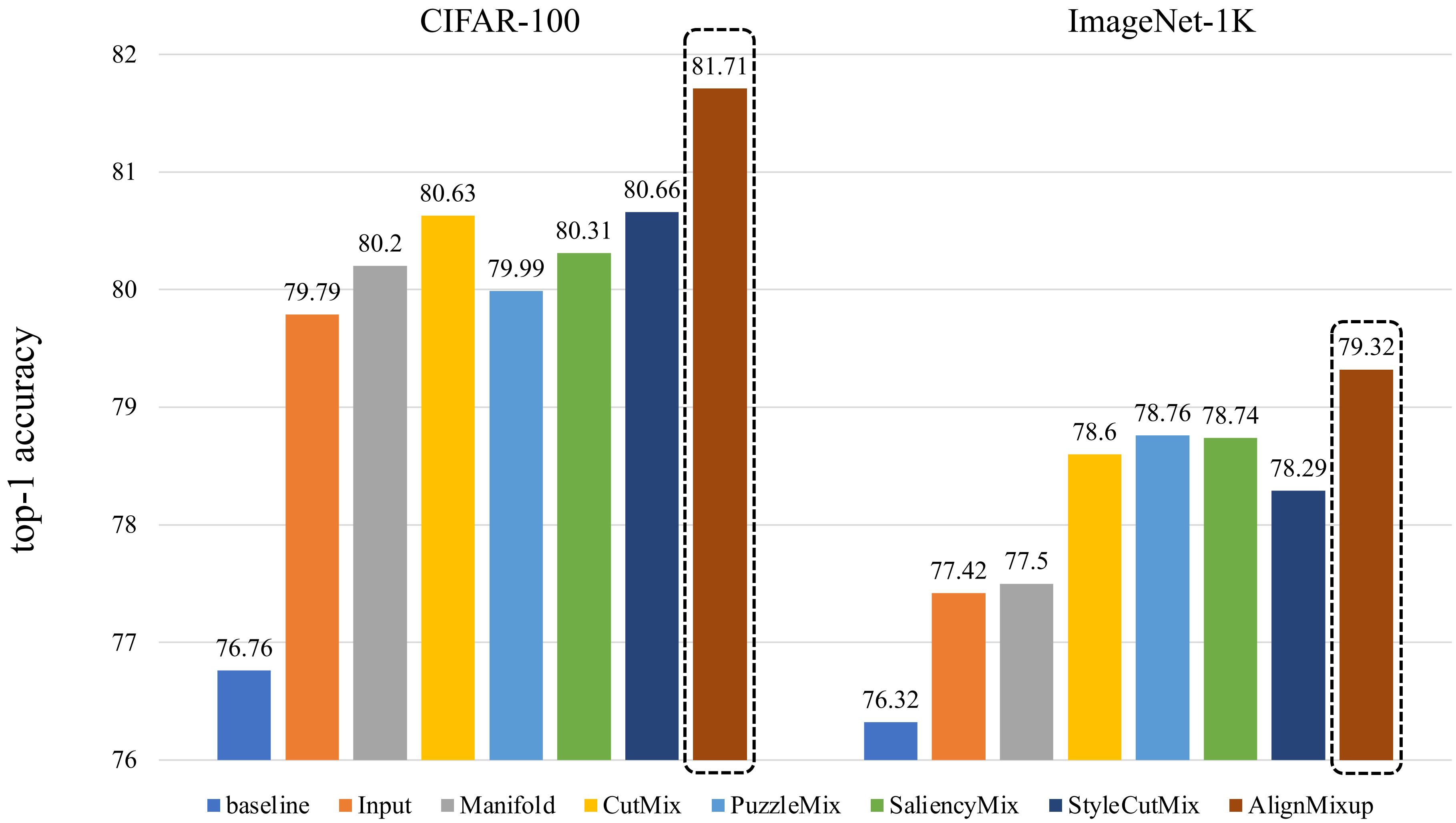
AlignMixup: visualizing alignment



$$\mathrm{mix}_\lambda\left(\mathbf{A}, \widetilde{\mathbf{A}}\right) = \lambda\mathbf{A} + (1-\lambda)\,\widetilde{\mathbf{A}}$$
$$\mathrm{mix}_\lambda\left(\mathbf{A}', \widetilde{\mathbf{A}}'\right) = \lambda\mathbf{A}' + (1-\lambda)\,\widetilde{\mathbf{A}}'$$

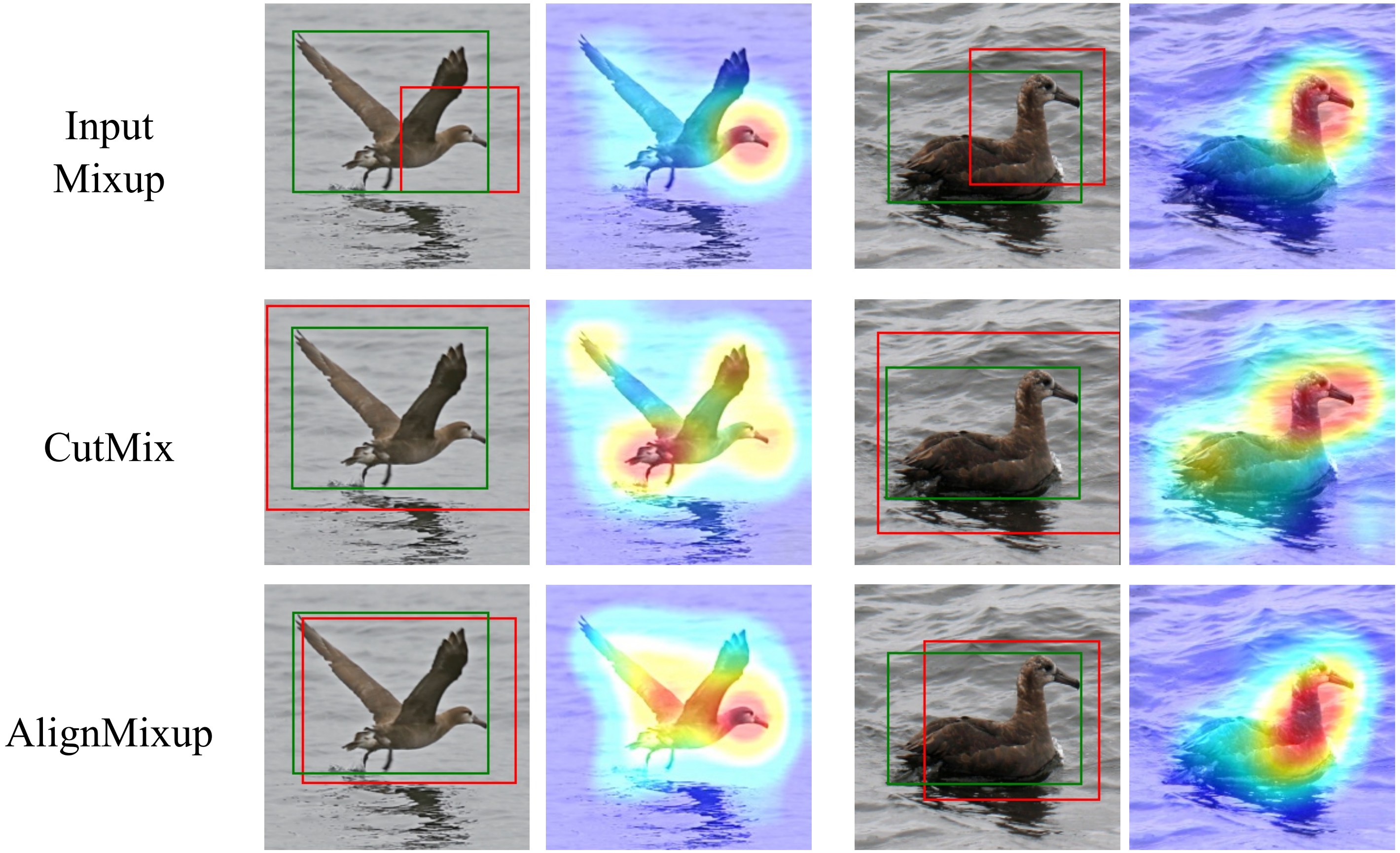
[images only for visualization; not used for training]

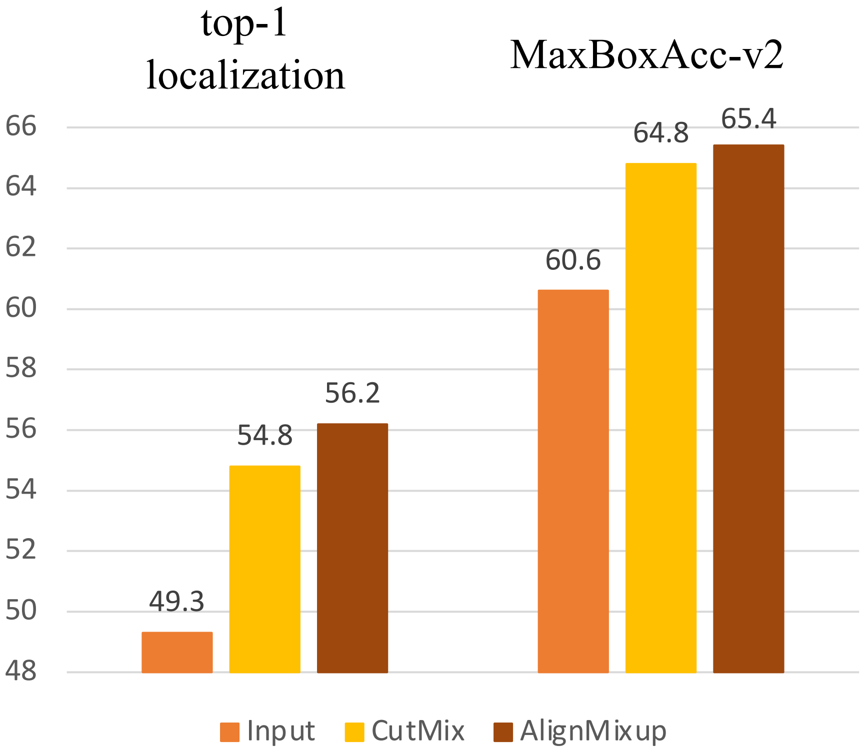
Image Classification



[additional results in the paper]

Weakly-Supervised Object Localization





[additional results in the paper]

See you on

24th June - Poster session 4.2!!