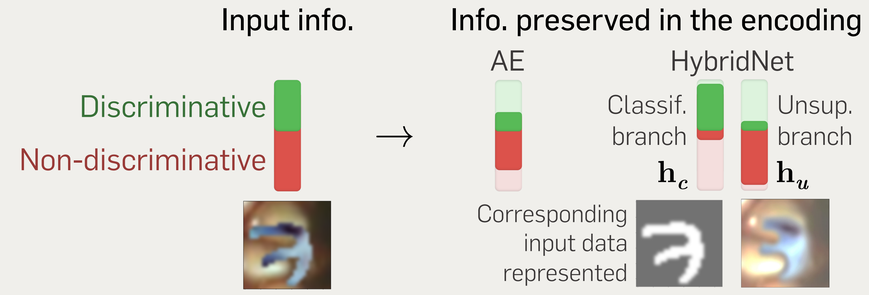
HYBRIDNET: CLASSIFICATION AND RECONSTRUCTION COOPERATION FOR SEMI-SUPERVISED LEARNING

Thomas Robert[1], Nicolas Thome[2], Matthieu Cord[1]
[1]Sorbonne Université (LIP6) & [2]CNAM – Paris, France

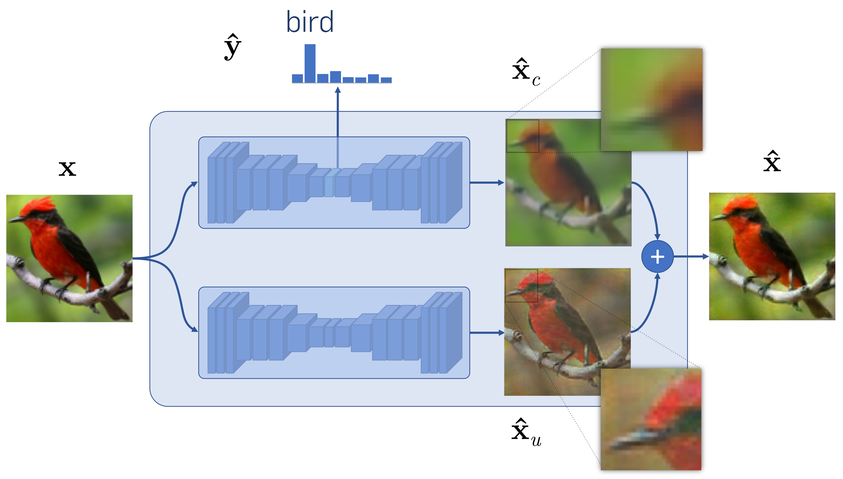
1. Context

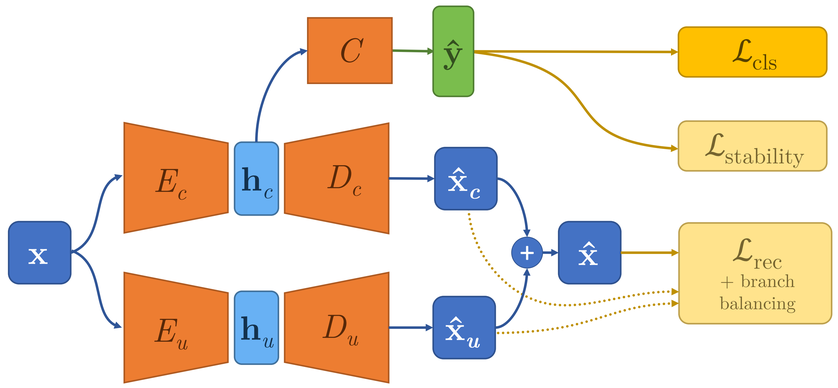
- **Improving classification** score on partially labeled data (< 5%)
- **Auto-encoder based semi-sup.** with classif. and reconstr.
- **Problem: antagonists costs**
  - *Classification* requires *information loss*
    ⇒ invariant features & good generalization
  - *Reconstruction* requires *information conservation*
    ⇒ correct reconstruction of the input
- **Related work:**
  - Auto-encoder w/ skip connections (LadderNet[1], SWWAE[2])
  - Classifier w/ stability based regularization (no reconstr.)[3,4,5]

2. Intuition & HybridNet architecture



- Explicitly separates discriminative and complementary information into two branches
- **Classification branch:** discriminative info. / partial reconstr.
- **Unsupervised branch:** complementary info. / complem. rec.





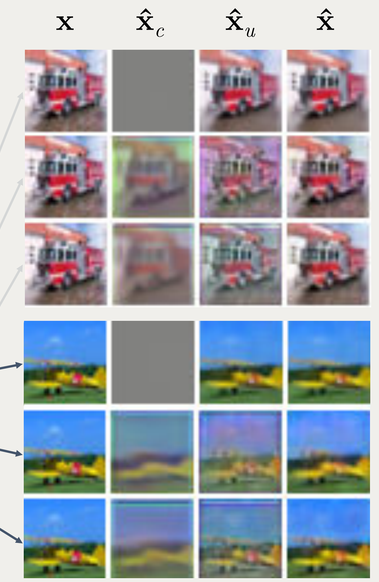
3. Controlling information separation

$$\mathcal{L} = \mathcal{L}_{\text{cls}} + \lambda_s \mathcal{L}_{\text{stability}} + \lambda_r \mathcal{L}_{\text{rec}}$$

- Controls the behavior of information separation
- Encourage discriminative / invariant features in $\boldsymbol{E_c}$
  - Classification loss $\quad \mathcal{L}_{\text{cls}} = \text{CrossEntropy}(\mathrm{y}, \hat{\mathrm{y}})$
  - Stability loss[3-5] $\quad \mathcal{L}_{\text{stability}} = \left\| \hat{\mathbf{y}}^{(k)} - \tilde{\mathbf{z}}^{(k)} \right\|$, $\tilde{\mathbf{z}}^{(k)} = \text{EMA}(\hat{\mathbf{y}}^{(k)})$
- Extract additional info. & balance branches
  - Reconstruction loss $\quad \mathcal{L}_{\text{rec}} = \|\hat{\mathrm{x}} - \mathrm{x}\|$
  - Branch balancing backpropagate $\mathcal{L}_{\text{rec}}$ only in the branch making the largest error btwn $\|\hat{\mathbf{x}}_c - \mathbf{x}\|$ and $\|\hat{\mathbf{x}}_u - \mathbf{x}\|$

Ablation study of the different loss terms

Classif.	Stability	Reconstr.	Branch balanc.	CIFAR-10 2000 labels
✓				71.5
✓	✓			74.6
✓		✓		72.4
✓		✓	✓	75.2
✓	✓	✓	✓	81.6



4. Results

Visualization of information separation

Shows loss of information in E_c and correction by E_u

Semi-sup. state-of-the-art results

Dataset	CIFAR-10		STL-10	SVHN
Nb. of labeled images	1000	4000	1000	1000
Nb. of unlabeled images	~50k		~100k	~600k
AE-based (Ladder[1])		20.40		
AE-based (SWWAE[2])			25.7	23.6
Stability only (MT[5])	10.10	6.23	16.8	4.2
Classif. baseline	45.22	15.45	18.0	10.0
HybridNet	**8.81**	**6.09**	**15.9**	**2.5**

- ResNet-based model. N labeled samples, $N/10$ per class, rest of the dataset unlabeled
- Surpasses comparable stability & AE-based baselines on standard benchmarks

5. Future & References

Conditional generative version through latent space manipulation with information disentanglement

[1] Rasmus *et al.* NIPS 15. SSL with ladder networks
[2] Zhao *et al.* ICLR Whp 16. Stacked What-Where Auto-encoders
[3] Sajjadi *et al.* NIPS 16. Regularization With Stochastic Transfo.
[4] Laine *et al.* ICLR 17. Temporal Ensembling for SSL
[5] Tarvainen *et al.* NIPS 17. Mean teachers are better role models