Knowledge Distillation and DCGAN

Thanh Nguyen

University Paris-Est, CNRS, ESIEE Paris, France

BusterAI - Dec 2020

Knowledge Distillation in Neural Network

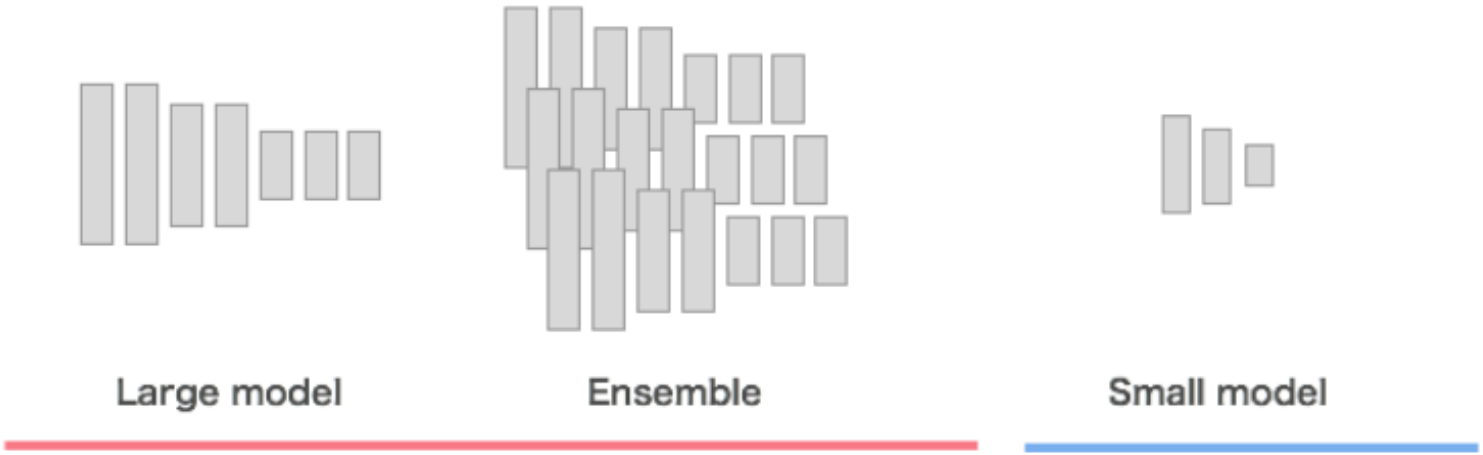
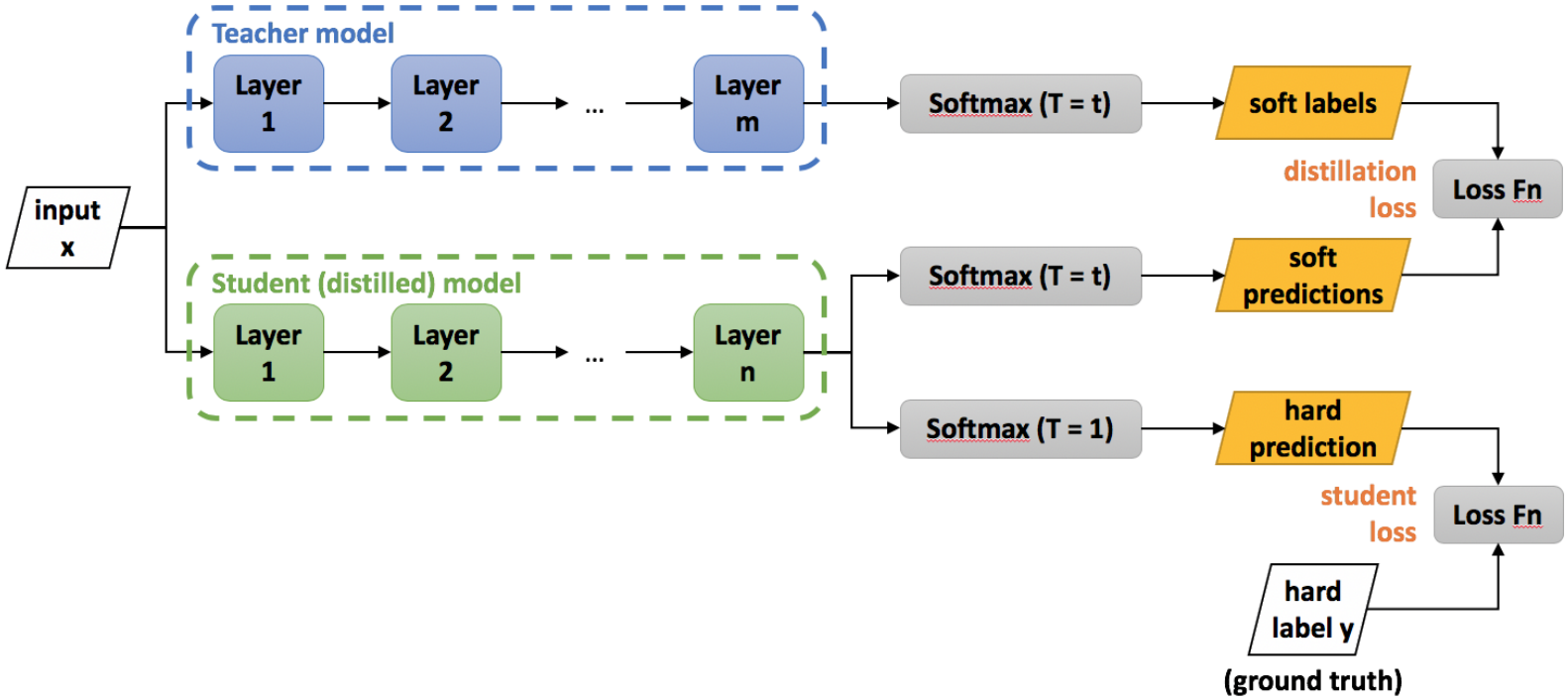


Figure: Models

Problem Statement

- To transfer cumbersome Teacher Model(s) into a lightweight Student Model.

Knowledge Distillation in Neural Network

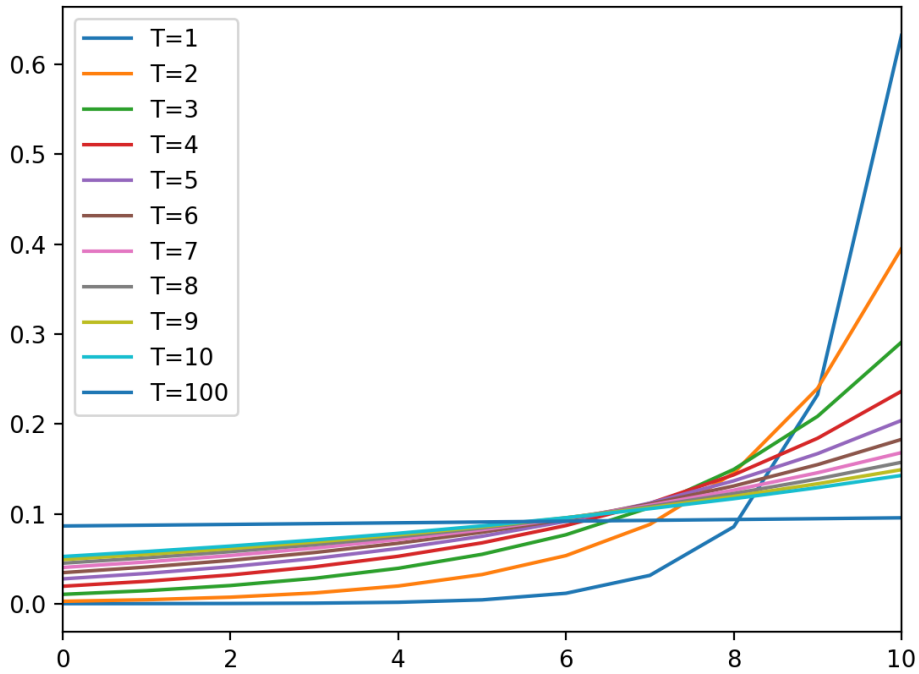


Approach To match **soft targets** of the Teacher(s) **tea** and the Student **std**.

- soft targets are the class probabilities produced by $\text{softmax}(X, T)$.
- transfer set: unlabeled data (X_1) and/or labeled data (X_2, Y_2).
- $loss_{distill} = CE(\text{softmax}(\text{std}(X_1), T > 1), \text{softmax}(\text{tea}(X_1), T > 1))$
- $loss_{student} = CE(\text{softmax}(\text{std}(X_2), T = 1), Y_2)$
- $loss = T^2 \cdot loss_{distill} + \alpha \cdot loss_{student}$

Knowledge Distillation in Neural Network

$$\text{softmax}(X, T) = \frac{\exp^{(X_i/T)}}{\sum_i \exp^{(X_i/T)}}$$



Approach To match **soft targets** of the Teacher(s) **tea** and the Student **std**.

- soft targets are the class probabilities produced by softmax of the logit X at temperature T.

Knowledge Distillation in Neural Network

Results

- Distilled model outperforms the same size model on the same training data.
- Training Specialists on huge dataset using Generalist.

Discussion

- Distillation soft targets as regularizers.
- The temperature T and entropy of the soft targets.
- Student model size and performance trade-off.

High-Resolution Neural Face Swapping

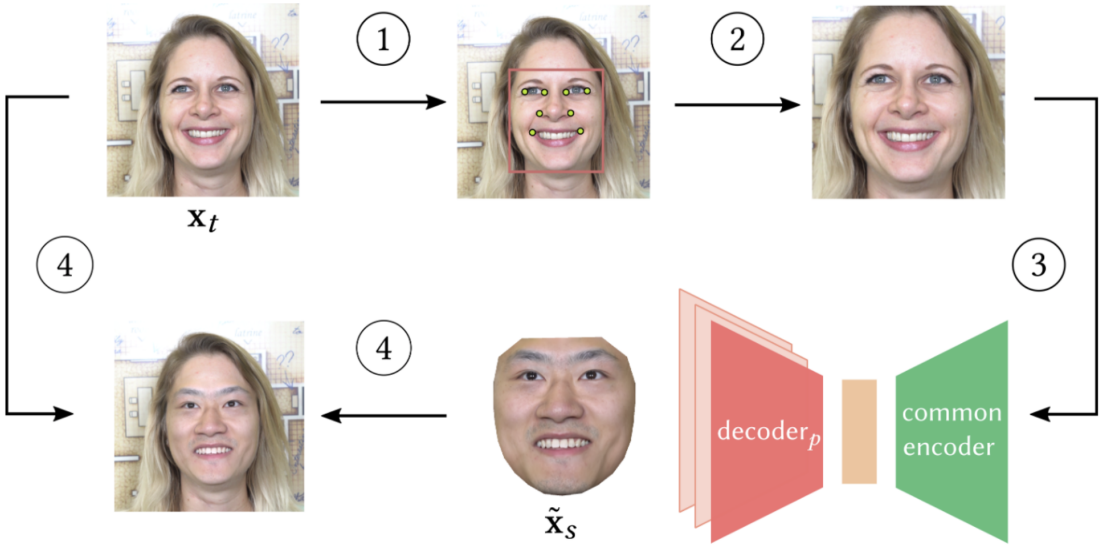


Figure: The face swap pipeline

Problem Statement

- To swap faces in image/video at megapixel resolution.

Approach To use DCGAN and blending.

- DCGAN: **progressive training** $\rightarrow$ high resolution output ($\approx$ StyleGAN).
- DCGAN: single encoder and multiple decoders $\rightarrow$ better generalize.
- blending: compositing, contrast preserve, landmark stabilization, etc.

High-Resolution Neural Face Swapping

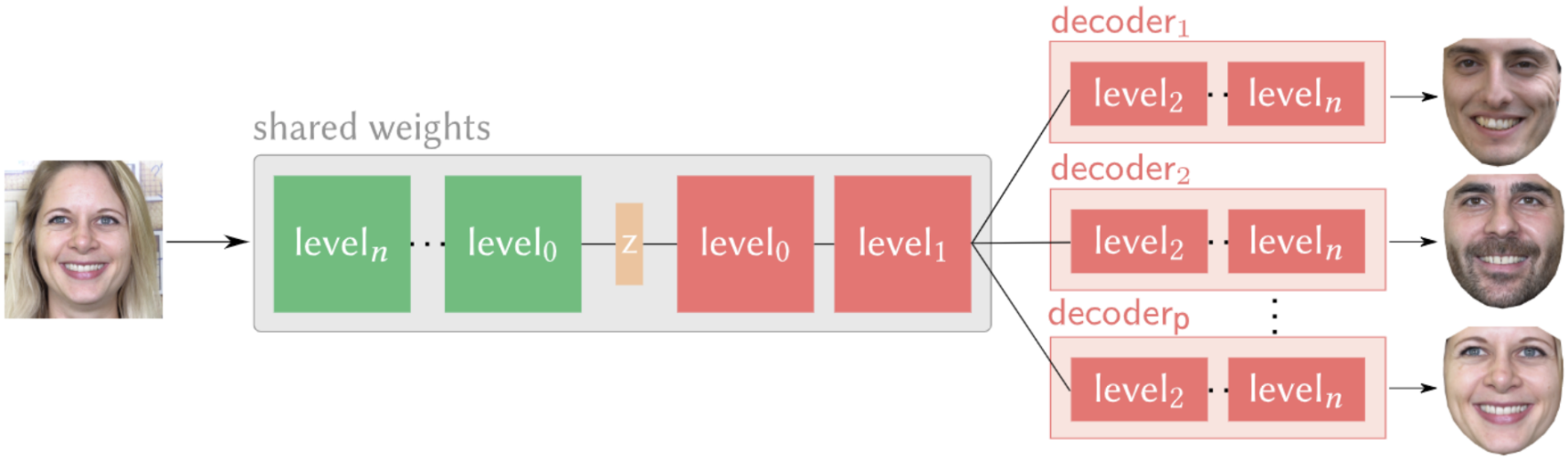


Figure: The multi comb DCGAN

Problem Statement

- To swap faces in image/video at megapixel resolution.

Approach To use DCGAN and blending.

- DCGAN: **progressive training** → high resolution output (≈ StyleGAN).
- DCGAN: single encoder and multiple decoders → better generalize.
- blending: compositing, contrast preserve, landmark stabilization, etc.

High-Resolution Neural Face Swapping

Results

- Higher resolutions, thanks to the StyleGAN.

Discussion

- The contribution is the pipeline.
- Blending, pre/post-processing play important roles.
- Blending regions are the most unrealistic parts $\rightarrow$ deepfake detector could pay more attention on those regions?

Thank you for your attention! Q&A