Quantifying Training Membership Information in the Hyperspherical Embedding Geometry of Face Recognition Models

Ünsal Öztürk[1] Sébastien Marcel[1,2]

[1]Idiap Research Institute, Martigny, Switzerland
[2]Université de Lausanne (UNIL), Lausanne, Switzerland

{unsal.ozturk, sebastien.marcel}@idiap.ch

Abstract

Face recognition models represent each face as an embedding vector on the unit hypersphere by clustering embeddings of the same identity while pushing different identities apart through angular-margin losses. Because these losses act only on training identities, non-member identities may form clusters with different geometric properties. In this paper, we quantify the magnitude of this difference and what training-time factors control it. We compute four statistics based on cluster geometry across 180 face recognition models in a factorial design over IResNet backbone size, loss head, training duration, and the number of training identities, and evaluate each configuration on nine benchmarks. Our results indicate that the number of training identities has the largest effect on member/non-member separability, while backbone and loss head contribute far less, and that, on a same-domain held-out reference, the geometric membership signal decreases monotonically as more identities are added to training. We provide an analysis of cross-domain (pose, age, quality, ethnicity) non-member benchmarks and report that these inflate the apparent membership signal. Finally, we fuse all four statistics with a learned classifier to reveal additional membership information beyond the best individual statistic.

1. Introduction

Modern face recognition (FR) systems are trained on large-scale datasets of face crops, typically collected from the internet [36]. To represent a person, the system takes multiple crops of that person's face and passes each through a deep neural network that produces a 512-dimensional embedding vector, normalised to lie on the unit hypersphere. The training of these models is primarily driven by angular-margin losses [10, 33, 26], which pull the embeddings of crops belonging to the same identity into a cluster on the hypersphere while pushing apart the clusters of different identities. During training, a classification head maps embeddings to identity logits and this head is discarded after training. When deployed, the system typically uses only the backbone neural network. Enrolled faces are stored as embeddings, and the similarity between two faces is measured by the inner product of their corresponding embeddings.

As angular-margin training directly optimises the compactness of each identity's cluster, we pose the following question: how much information about training-set membership is retained in the geometry of these embedding clusters? If training identities form systematically tighter or better-separated clusters than unseen identities, then the embedding geometry carries a residual membership signal.

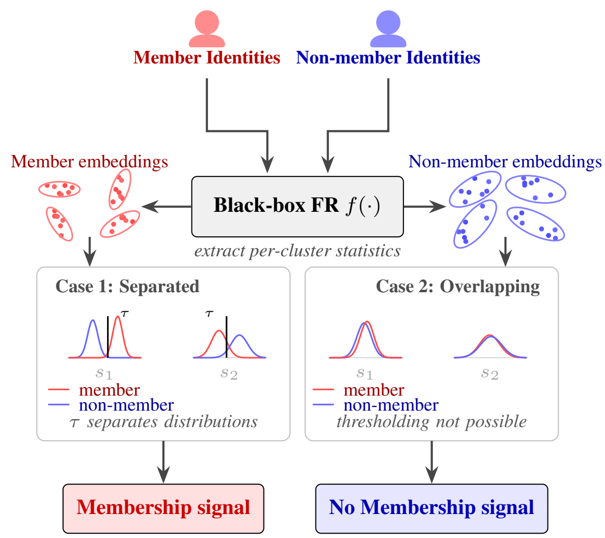


Figure 1. Measurement pipeline. Per-identity statistics $(s_1, s_2, \dots)$ are computed from the embedding clusters of a face recognition model. If the distributions of these statistics differ between member identities (present in the training set) and non-member identities (absent from it) as shown on the left, the explicit cluster geometry carries membership information on these statistics; if they coincide (right), it does not.

Project page: https://idiap.ch/paper/quantifying-membership-inference.

Quantifying this residual signal is related to membership inference attacks (MIA), a line of work investigating whether a given sample or identity was part of a model's training data [31, 5]. In practice, membership inference serves as an auditing tool for data privacy by allowing individuals or regulators to verify whether a model was trained on data belonging to a specific person in the context of data protection rights such as the right to erasure under the GDPR [14]. MIA literature primarily focuses on closed-set classifiers that expose prediction confidences, loss values, or the model weights directly [31, 29, 5, 18]. In our setting, we consider open-set face recognition where neither logits nor losses nor model weights are available, so the membership signal must be sought entirely in the geometry of the embeddings themselves.

To measure this signal, we compute four different statistics capturing different aspects of cluster geometry on the hypersphere and measure how much they differ between training and unseen identities (Figure 1). We train 180 face recognition models from scratch (Section 4.2) and evaluate each model on nine benchmarks, including a same-domain held-out split that controls for the pose, age, quality, and ethnicity shifts present in the other benchmarks. To our knowledge, this constitutes the largest systematic study of membership information in face embedding geometry. Our contributions are: **(1)** we quantify the membership information retained in four geometric statistics computed from face embeddings; **(2)** we identify the primary factors that control how much of this information persists; **(3)** we show that cross-domain non-member benchmarks (e.g. differing in pose or age) inflate the apparent membership signal by conflating domain shift with training-attributable differences; **(4)** we demonstrate that fusing all statistics captures membership information beyond the best individual statistic.

2. Related Work

Membership inference on classifiers. Membership inference literature targets closed-set classifiers, where the attacker can observe prediction confidences, logits, or loss values of the target model. In the open-set face recognition setting, however, none of these signals are available and this has motivated a line of work operating directly on embeddings. Shokri *et al.* [31] introduce the shadow-model paradigm and train substitute classifiers to learn a binary meta-classifier on output distributions. Salem *et al.* [29] relax many of these assumptions and show that a single shadow model and only the top-k predictions often suffice. Carlini *et al.* [5] recast MIA as a likelihood-ratio attack (LiRA) and derive near-optimal attacks. They show that per-example hardness varies by orders of magnitude. Hu *et al.* [18] provide a comprehensive survey.

Membership inference on embedding models. Existing embedding-based approaches differ in the granularity of the membership question: some ask whether a specific image was used for training (sample level), while others ask whether *any* image of a given person was included (identity level). Li *et al.* [22] formalise identity-level inference on metric-embedding models and use intra-similarity and inter-dissimilarity of probe embeddings as attack features on person re-identification and face recognition. Chen *et al.* [7] propose FACE-AUDITOR, a system that queries few-shot face recognition models with designed probing sets and uses the returned similarity scores as auditing features. They report accuracies up to 99%. SLMIA-SR [6] addresses the analogous problem in speaker recognition by training shadow models with controlled membership mixing ratios. All three operate at the identity level. EncoderMI [23], by contrast, operates at the sample level on contrastive-learning encoders including CLIP and exploits the observation that an overfitted encoder produces more similar representations for augmented views of a training sample than for an unseen one.

Table 1. **Comparison with prior membership inference studies.** Level: identity-level or sample-level membership question. # Mod.: number of target models evaluated. Fact.: whether a factorial experimental design was used.

Work	Domain	Level	# Mod.	Fact.
Embedding-level (black-box)				
Li *et al.* [22]	FR / re-identification	Identity	6	✗
FACE-AUD. [7]	FR / few-shot	Identity	4	✗
SLMIA-SR [6]	Speaker recognition	Identity	5	✗
EncoderMI [23]	Contrastive learning	Sample	4	✗
Model-internals (white-box)				
MINT [9]	FR	Sample	3	✗
DeAlcala [8]	FR	Sample	9	✗*
Mancera [24]	Object classification	Sample	4	✗
Huang [20]	FR	Sample	2	✗
Ours	**FR**	**Identity**	**180**	✓

* Varies factors individually but not in a fully crossed design.

White-box approaches on face recognition. A parallel line of work allows access to model internals. The motivation is that intermediate representations such as activation maps, batch-normalisation statistics, or gradient signals may carry a stronger membership trace than the final embedding alone, at the cost of requiring more privileged access to the model. These methods also typically operate at the sample level, determining whether a specific image (rather than a person) was included in training. The Membership Inference Test (MINT) of DeAlcala *et al.* [9] trains an MLP or CNN on intermediate activation maps to classify individual face images as member or non-member and achieves up to 90% accuracy on three ResNet-100 models. A companion study [8] analyses factors that affect MINT performance, including loss function, dropout, and number of training epochs. Mancera *et al.* [24] extend MINT to object classification and report 70–80% precision depending on layer depth. Huang *et al.* [20] exploit distances between intermediate features and batch-normalisation parameters for sample-level membership and model-inversion attacks [13]. Even with this level of access, Rezaei and Liu [28] show that the advantage over random guessing can be modest once the model generalises well.

Memorisation and embedding geometry. The question of why a geometric membership signal might exist at all connects to a discussion on memorisation in deep networks. If a model must memorise certain training examples to achieve low error, particularly on the tail of the data distribution, then its internal representations and output embeddings may carry traces of those examples. Feldman [12] formalises this argument, showing theoretically that long-tailed distributions necessitate memorisation for near-optimal generalisation. Angular-margin losses [10, 33, 26] amplify this effect, potentially making training identities geometrically distinguishable from unseen ones. Unlearning methods [4, 15] and differential privacy [1] aim to erase or bound such signals.

Comparison with this work. Our contribution is a controlled measurement study of how much membership information geometric statistics carry, not a stronger attack feature. Table 1 summarises the literature. Prior embedding-level work proposes specific attack features and demonstrates that membership inference is feasible through evaluating a small number of models trained under a single configuration. Our study differs in three respects. First, the goal: we measure how much membership information is retained in the explicit cluster geometry and identify which training factors control the amount that persists. Second, the experimental scale: 180 models, whereas the largest prior study evaluates nine. Third, domain control: we add a same-domain held-out reference, which aims to separate training-attributable signal from domain-shift effects. Several of the per-identity statistics we evaluate have already been used as features in prior attacks. For instance, the intra-similarity score proposed by Li *et al.* [22] is equivalent to the mean pairwise cosine similarity that we include as a baseline statistic (defined in Section 3.2).

3. Methodology

3.1. Access Model and Measurement Procedure

We assume that an auditor can query an FR model f with $k \geq 2$ probe images of a target identity y and obtain the corresponding ℓ_2-normalised embeddings. From the resulting cluster, the auditor computes four per-identity statistics (Section 3.2) and uses them to assess if the cluster geometry carries information about y's membership in the training set. For each model we train, we compute these statistics on member identities (from WebFace4M) and on non-member identities drawn from the benchmarks in Table 2.

For each statistic, we obtain a scalar score per identity. To quantify how much membership information a statistic carries, we sweep a decision threshold τ across the range of observed values and classify an identity as a member if its score exceeds τ (or falls below τ for statistics where lower values indicate membership). We report the area under the resulting ROC curve (AUC), hereafter threshold-classifier AUC, as a threshold-free summary, which is equal to the probability that a randomly chosen member scores higher than a randomly chosen non-member. This serves as our primary measure of the membership information carried by a given statistic. We use AUC rather than the true positive rate at a fixed false positive rate, because our framing concerns separability across all decision thresholds rather than the operating point most relevant for a deployable attack. Of the four statistics defined in the next section, two require only the embeddings and two additionally require the training-time scale and margin hyperparameters. We refer to these two access levels as black-box (embeddings only) and grey-box (embeddings plus s and m). The grey-box setting is realistic when models are released alongside their training configurations.

3.2. Per-Identity Geometric Statistics

Given an identity y with $k \geq 2$ probe images, we extract embeddings $\{e_1, \ldots, e_k\}$ via the target model and compute four statistics from the resulting cluster. Each statistic captures a different aspect of cluster shape. The underlying intuition is that an angular-margin loss explicitly optimises the representation of training identities, so members are expected to form tighter clusters on the embedding hypersphere than non-members. If that expectation holds, the statistic will carry membership information; the question is how much, and what factors control it.

Pairwise cosine similarity (PAIRCOS). The mean cosine similarity over pairs of embeddings within the identity:

$$\bar{s} = \binom{k}{2}^{-1} \sum_{i<j} \cos(e_i, e_j).$$

Higher values indicate a more coherent cluster.

vMF concentration κ. We model the per-identity embedding distribution as a von Mises-Fisher (vMF) distribution on the unit hypersphere and estimate the concentration parameter κ via maximum likelihood. We use the closed-form approximation of Banerjee *et al.* [3]:

$$\hat{\kappa}_0 = \frac{\bar{R}\,(d - \bar{R}^2)}{1 - \bar{R}^2}, \qquad \bar{R} = \left\| \frac{1}{k} \sum_{i=1}^{k} e_i \right\|,$$

where $d = 512$ is the embedding dimensionality and $\bar{R}$ is the mean resultant length. Higher $\hat{\kappa}$ corresponds to greater directional concentration. Note that ordering identities by $\hat{\kappa}$ is equivalent to ordering them by cluster inertia

Table 2. **Dataset overview for membership evaluation.** "Eval. subj.": identities with $\geq$2 crops (required for per-identity cluster statistics). †For WebFace4M this is the held-out split we sample for evaluation; the dataset itself contains far more identities with $\geq$2 crops.

Dataset	Primary variation	Subjects	Eval. subj.	Images	Covariate metadata
Cross-distribution benchmarks (non-member identities)					
LFW [19]	Unconstrained (baseline)	5,749	1,680	13,233	— (subject name only)
CPLFW [35]	Cross-pose (subset of LFW identities)	3,929	3,909	11,652	yaw (estimated from landmarks)
CFP-FF [30]	Frontal-only probes	500	500	7,000	pose $\in$ {frontal}
CFP-FP [30]	Frontal+profile probes		500	7,000	pose $\in$ {frontal, profile}
AgeDB-30 [27]	Large intra-identity age gap ($\geq$30 yr)	440	440	12,240	age (3–100 yr), gender
XQLFW [21]	Low image quality (paired with LFW)	5,749	1,672	13,233	— (quality implicit by design)
RFW [34]	Ethnicity-balanced (4 groups)	11,429	11,429	40,607	ethnicity, gender, nationality
Large-scale benchmarks					
IJB-C [25]	Unconstrained, mixed image+video	3,531	3,531	469,376	ethnicity, gender, nationality
Training data and within-distribution evaluation					
WebFace4M [36]	Large-scale training source	205,990	$\leq$ 5,000†	4,000,000	— (subject ID only)

$I = k^{-1} \sum_i \|e_i - \mu\|^2$, where $\mu = k^{-1} \sum_i e_i$ is the cluster mean. For ℓ_2-normalised embeddings, I reduces to $1 - \bar{R}^2$, and $\hat{\kappa}_0$ is strictly decreasing in $1 - \bar{R}^2$, so the two quantities produce identical identity rankings.

Penalised logit (PenLogit). To approximate the margin-penalised logit that the training loss would compute for the cluster, we use a leave-one-out scheme. For each embedding e_i, we compute a prototype $\hat{w}_{\setminus i} = \bar{e}_{\setminus i}/\|\bar{e}_{\setminus i}\|$ as the ℓ_2-normalised mean of the remaining $k-1$ embeddings, and the margin-penalised logit is

$$\ell_i = s \cdot g\left(\cos\angle\left(e_i, \hat{w}_{\setminus i}\right),\ m\right),$$

where s and m are the scale and margin of the loss head, and g is the head-specific penalisation function (e.g. $g(\cos\theta, m) = \cos(\theta + m)$ for ArcFace). The statistic value is the mean $\bar{\ell} = k^{-1} \sum_i \ell_i$. A higher penalised logit indicates that the embeddings are tightly aligned to their own prototype.

Prototype softmax CE (ProtoCE). We construct a surrogate of the training loss. For each identity y, we compute a class prototype $\hat{w}_y = \mu/\|\mu\|$ and treat all other identities in the evaluation set as negative classes. The cross-entropy loss with the margin-modified logit for the target class and standard cosine logits for the negatives gives

$$\mathcal{L}_y = -\log \frac{\exp\big(s \cdot g(\cos\theta_y,\, m)\big)}{\exp\big(s \cdot g(\cos\theta_y,\, m)\big) + \sum_{j \neq y} \exp(s \cdot \cos\theta_j)}$$

We report $-\mathcal{L}_y$ so that higher values indicate membership, consistent with the sign convention of the other statistics. Because the softmax denominator depends on the composition and size of the evaluation set, ProtoCE values are not directly comparable across benchmarks with different identity pools. Throughout, all AUC values are oriented so that AUC > 0.5 indicates better-than-chance separation.

4. Experimental Setup

4.1. Training Data

All models are trained on WebFace4M [36], a large-scale face dataset containing approximately 4 million images of 205,990 identities. WebFace4M is well-suited to this study because its size permits creating training subsets that span several orders of magnitude in identity count while still providing sufficient images per identity for meaningful embedding statistics. Every probe image is aligned to a canonical 112×112 crop with the standard InsightFace five-landmark pipeline.

4.2. Ablation Design

We vary four factors in a fully crossed design.

Backbone capacity. We train three IResNet architectures: IResNet-34, IResNet-50, and IResNet-100 [16, 10].

Angular-margin loss head. We pair each backbone with three angular-margin loss heads: ArcFace [10] ($s = 64$, $m = 0.5$), CosFace [33] ($s = 64$, $m = 0.4$), and MagFace [26] ($s = 64$, per-sample $m \in [0.45, 0.8]$ depending on the un-normalised feature magnitude).

Training-set size (n_{ids}). To study the effect of identity coverage, we create five partitions by randomly sampling $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}\}$ identities (with a fixed seed), plus one partition that uses all available identities. The identities not selected for training serve as same-domain non-members in the held-out WebFace4M benchmark.

Training duration. We train the backbones for 5, 10, 20, and 40 epochs to track the evolution of cluster-statistic separability. Longer training exposes each identity to more gradient updates and may therefore affect the gap between member and non-member statistic distributions.

The full grid spans 3 backbones $\times$ 3 losses $\times$ 5 n_{ids} levels $\times$ 4 epochs $=$ 180 model configurations. All models are trained with SGD (momentum 0.9, weight decay 5×10^{-4}), a polynomial-decay learning-rate schedule starting at 0.1 with one warmup epoch, batch size 128, mixed-precision (FP16), and random horizontal flips.

4.3. Evaluation Benchmarks

The nine benchmarks (Table 2) supply standard verification protocols for measuring model quality and provide non-member pools with diverse domain characteristics so that we can study how pose, quality, and demographics affect the apparent cluster-statistic separability.

Eight of the nine benchmarks are drawn from sources disjoint from WebFace4M (we verified that no identity overlap exists by cross-checking subject identifiers); Table 2 lists their sizes and available covariate metadata. Two dataset pairs form natural experiments (matched comparisons where a single attribute varies). *CFP-FF* vs. *CFP-FP* [30] contrasts frontal-only against frontal-plus-profile probes for the same 500 identities, so any statistic difference is attributable to pose. *LFW* [19] vs. *XQLFW* [21] shares the same identity set and pair protocol but degrades one image per pair to isolate image quality. *CPLFW* [35] (cross-pose LFW), *AgeDB-30* [27] (age gap ≥30 yr), *RFW* [34] (ethnicity-balanced, four groups), and *IJB-C* [25] complete the cross-domain set.

Held-out WebFace4M split. The ninth benchmark is constructed from WebFace4M itself: for each model trained on n_{ids} identities, we sample $\min(n_{\text{ids}}, 2{,}500)$ members and 2,500 non-members from the remaining subjects; at n_{ids}=all no non-members remain. As both groups share the same acquisition pipeline, AUCs are arguably attributable to the cluster-geometry difference between seen and unseen identities rather than to domain shift.

5. Results and Discussion

This section presents experimental results and analysis for the 180-model grid described in Section 4. We examine verification performance, the effect of training-set size and duration on membership separability, distributional confounds, and statistic fusion. Full per-epoch, per-model results and ROC/DET curves are provided in the supplementary material.

5.1. Verification Performance

We first establish that our model grid spans a meaningful range of verification quality. Table 4 reports the verification equal error rate (EER) at epoch 40 for every combination of backbone, loss head, and n_{ids} on the cross-domain benchmarks. The table is dense and provided primarily as a per-cell reference, and the discussion below focuses on patterns and points back to specific cells only as needed. Performance generally improves with n_{ids}, with diminishing gains beyond 100K identities. Once n_{ids} is held fixed, backbone depth provides small reductions in verification EER on domain-shifted benchmarks, and the choice of loss head

Table 3. **What controls membership information in cluster geometry? Top**: partial $\hat{\eta}^2_p$ from a Type-II ANOVA (each factor tested after adjusting for all others) on the AUC of a threshold classifier separating member from non-member identities ($N = 180$ models); values close to 1 indicate that the factor explains nearly all variance in this separability. **Bottom**: distributional-shift effect sizes (Cohen's d', Spearman ρ, Kruskal–Wallis ϵ^2) averaged over all models; larger absolute values indicate a stronger confound.

	PairCos	κ	*PenLogit*	*ProtoCE*
Model-level ANOVA (partial $\hat{\eta}^2_p$)				
n_{ids}	.99	.95	.99	.95
Epoch	.99	.89	.98	.92
Backbone	.09	.01	.02	.04
Loss	.04	.01	.01	.01
$n_{\text{ids}}\times$Ep	.98	.92	.97	.92
Residual	.00	.03	.01	.02
Marginal means (PairCos AUC)				
n_{ids}: 1K 0.90 · 10K 0.81 · 50K 0.73 · 100K 0.71 · all 0.71				
Epoch: ep5 0.69 · ep10 0.75 · ep20 0.79 · ep40 0.85				
Distributional shift (averaged over all 180 models)				
Cohen's d'				
XQLFW − LFW	−1.64	−0.48	−1.81	+1.13
CFP-FP − CFP-FF	−2.27	−1.13	−2.98	+1.14
CPLFW − LFW	−1.72	−0.72	−2.06	+0.31
Spearman ρ				
AgeDB (age σ)	+0.01	−0.26	+0.17	+0.33
Kruskal-Wallis ϵ^2				
RFW (4 groups)	0.017	0.017	0.018	0.123

has an even smaller effect. Notably, the highest cluster-statistic separability occurs at n_{ids}=1K, where verification performance is worst. As n_{ids} grows, verification improves while separability decreases. However, non-trivial separability persists even at n_{ids} levels that achieve competitive verification EER, so the patterns reported below are not confined to poorly performing models.

Table 4 also reports the threshold-classifier AUC for all four cluster statistics. No single statistic achieves the highest AUC across all benchmarks. PenLogit ranks first most often at moderate-to-high n_{ids}, in particular on the pose- and quality-degraded sets, whereas ProtoCE attains the highest AUC on the frontal and high-quality sets (CFP-FF, LFW, IJB-C), so the best statistic remains dataset-dependent. Because the eight non-WebFace4M benchmarks differ from the training source in pose, quality, and demographics, the observed AUC conflates training-attributable and domain-induced differences, which we discuss in Section 5.3. All four statistics exhibit similar qualitative trends across factors, so we show only PairCos distributions in Figure 2.

5.2. Effect of n_{ids} and Training Duration

Figure 2 presents the main finding of this study. Panel (c) shows that the PairCos distributions of members and non-members converge as n_{ids} increases from 1K to all identities. At n_{ids}=1K the distributions are well separated, whereas at n_{ids}=100K and beyond the separation narrows,

Table 4. **Verification performance and membership separability at epoch 40.** **E** = EER, **C** = PairCos, **κ** = vMF-κ, **L** = PenLogit, **S** = ProtoCE AUC. Green = low EER / high AUC, red = high EER / low AUC; cyan = n_{ids}=all AUC $\geq$ 0.8; **bold** = best per backbone. Lower EER means better recognition, and higher AUC means more membership leakage.

Backbone Head		LFW					XQLFW					CPLFW					RFW					CFP-FP					CFP-FF					AgeDB					IJB-C				
		1K	10K	50K	100K	All	1K	10K	50K	100K	All	1K	10K	50K	100K	All	1K	10K	50K	100K	All	1K	10K	50K	100K	All	1K	10K	50K	100K	All	1K	10K	50K	100K	All	1K	10K	50K	100K	All
IResNet34 *ArcFace*	*E*	**7.7**	1.3	0.4	**0.2**	0.3	**21.3**	**12.8**	**7.3**	5.8	5.3	39.6	20.7	10.7	9.0	8.4	34.6	17.4	6.3	**3.9**	2.8	**26.3**	**7.7**	**1.8**	**1.1**	1.0	**9.6**	1.4	0.3	0.3	**0.1**	**27.1**	**9.6**	3.6	3.1	2.8	**8.3**	2.9	1.7	1.5	1.3
	C	.99	**.92**	.56	.71	**.77**	.99	**1.**	**.96**	.91	.89	**1.**	**1.**	**.90**	.80	.75	1.	**.99**	**.78**	.61	.52	**1.**	**1.**	**.94**	.83	.76	**1.**	**.99**	.53	.71	.78	**1.**	1.	**.96**	.84	.75	**.99**	**.94**	**.72**	.62	**.58**
	κ	**.78**	**.60**	.84	.89	.90	**.83**	**.94**	**.70**	.58	.54	**.96**	**.90**	.61	.71	.74	**.98**	**.93**	.52	.67	.72	**1.**	**1.**	**.82**	.63	.55	**1.**	**.99**	.51	.68	.73	**1.**	.99	**.92**	.81	.73	**.99**	**.96**	**.78**	.69	**.65**
	L	1.	**.98**	**.74**	**.59**	**.51**	1.	**1.**	.99	.97	.96	**1.**	**1.**	**.99**	**.97**	.95	1.	1.	**.90**	.79	.73	**1.**	**1.**	**.97**	.91	.87	**1.**	**.96**	.54	.72	.79	**1.**	1.	**.93**	.81	.73	**.96**	**.90**	**.63**	.53	.52
	S	.78	.59	.86	.90	.90	.93	.93	.63	.54	.55	.83	.82	.78	.83	.83	**.99**	.97	.56	**.63**	**.66**	1.	**1.**	.65	.58	.61	1.	**.98**	.55	.72	.73	.99	.98	.85	.70	.66	**1.**	**1.**	.93	.84	.83
IResNet34 *CosFace*	*E*	9.3	1.3	**0.4**	0.2	0.3	22.8	13.3	7.5	**5.7**	5.5	**39.4**	21.0	**10.6**	9.0	**8.1**	**34.2**	17.6	**6.1**	4.0	**2.6**	27.8	8.7	1.8	1.2	0.9	10.2	**1.2**	**0.3**	0.3	0.1	27.2	10.4	3.8	3.0	**2.6**	8.7	3.1	1.7	1.5	1.4
	C	.97	.89	**.58**	.71	.77	.97	.99	.94	.91	.88	1.	1.	.89	.80	**.75**	1.	.98	.77	**.61**	**.53**	**1.**	1.	.93	.83	.75	1.	.96	**.58**	**.72**	.78	**1.**	**1.**	.96	**.85**	**.76**	.97	.92	.69	.60	.56
	κ	.62	.51	**.86**	**.89**	**.91**	.65	.88	.61	.53	.50	.83	.80	**.69**	**.76**	**.78**	.94	.89	**.57**	**.69**	**.72**	1.	.99	.77	.59	.52	1.	.96	**.55**	.69	.73	.99	.99	.92	**.81**	**.74**	.98	.94	.76	.68	.64
	L	.99	.97	.71	.59	.51	1.	1.	.99	.97	.96	1.	1.	.99	.97	**.95**	1.	.99	.89	**.80**	**.73**	1.	1.	.96	.91	.87	.99	.92	**.59**	**.73**	**.79**	1.	1.	.92	**.82**	**.73**	.93	.87	.60	.52	**.53**
	S	.67	.53	**.87**	**.91**	.90	.93	.90	.62	.54	.56	.57	.67	**.83**	**.86**	**.84**	.97	.95	.56	.62	.65	.99	.99	.57	**.62**	**.62**	1.	.96	.59	.73	.73	.98	.97	.84	.69	.65	.99	1.	**.94**	**.87**	**.85**
IResNet34 *MagFace*	*E*	8.9	**1.2**	0.5	0.4	**0.2**	23.0	13.8	7.5	5.9	**5.1**	40.1	**20.6**	10.7	**8.9**	8.3	34.3	**17.4**	6.2	4.0	2.8	28.5	7.9	1.9	1.2	**0.9**	10.5	1.4	0.3	**0.2**	0.2	27.4	10.2	**3.6**	**2.9**	2.8	8.7	**2.9**	**1.6**	**1.4**	**1.3**
	C	**.99**	.91	.58	**.72**	.77	**.99**	1.	.96	**.91**	**.89**	1.	1.	.89	**.81**	.75	**1.**	.99	.77	.60	.52	**1.**	1.	.94	**.84**	**.76**	1.	.98	.56	.72	**.78**	**1.**	1.	.96	.84	.75	.99	.94	.71	**.62**	.57
	κ	.78	.57	.84	.89	.90	.80	.93	.69	**.59**	**.54**	.95	.88	.62	.71	.74	.98	.93	.53	.67	.72	1.	1.	.82	**.63**	**.55**	**1.**	.98	.53	**.69**	**.74**	1.	**.99**	.92	.81	.73	.99	.96	.78	**.69**	.65
	L	**1.**	.98	.72	.59	.51	**1.**	1.	**.99**	**.97**	**.96**	1.	1.	.99	.97	.95	**1.**	**1.**	.89	.79	.72	1.	1.	.97	**.92**	**.87**	1.	.95	.57	.73	.79	**1.**	**1.**	.92	.81	.73	.96	.89	.62	**.53**	.52
	S	**.79**	**.60**	.86	.90	**.90**	**.93**	**.95**	**.72**	**.63**	**.64**	**.88**	**.86**	.70	.78	.79	.99	**.97**	**.64**	.55	.59	**1.**	1.	**.72**	.50	.53	**1.**	.97	**.62**	**.76**	**.75**	**.99**	**.99**	**.86**	**.73**	**.70**	.99	.99	.87	.79	.78
IResNet50 *ArcFace*	*E*	**8.2**	1.2	0.4	**0.2**	0.2	**21.0**	**12.5**	7.0	**5.2**	4.5	**39.3**	20.6	10.0	**8.3**	7.7	34.3	17.5	6.0	3.5	2.1	**26.0**	7.9	1.6	0.9	**0.6**	**9.5**	1.5	**0.3**	0.1	0.2	**26.4**	9.5	3.9	3.0	**2.5**	**8.2**	3.0	1.5	1.4	**1.2**
	C	.99	**.92**	.50	.69	.76	.99	**1.**	**.96**	**.92**	.89	**1.**	**1.**	**.92**	.82	.75	1.	**.99**	**.83**	**.64**	.53	**1.**	**1.**	**.96**	**.86**	.77	1.	**.99**	**.55**	.67	.76	**1.**	1.	**.98**	.88	.78	.99	**.95**	**.75**	**.65**	.59
	κ	.80	**.62**	.81	.88	.90	**.86**	**.95**	**.74**	**.62**	.57	**.98**	**.90**	.57	.70	.74	**.98**	**.94**	**.56**	.64	.71	**1.**	**1.**	**.88**	**.67**	.56	**1.**	**.99**	**.58**	.64	.72	1.	**1.**	**.95**	.85	.76	.99	**.97**	**.81**	**.71**	.66
	L	1.	**.98**	**.77**	**.62**	**.53**	1.	1.	**.99**	**.98**	.97	1.	1.	**.99**	.97	.95	1.	**1.**	**.92**	**.82**	.74	**1.**	**1.**	**.98**	.93	.88	1.	**.97**	**.53**	.67	.77	**1.**	**1.**	**.95**	.85	.75	.97	**.91**	**.67**	**.56**	.50
	S	.80	.61	.84	.89	.89	**.94**	.94	.67	.57	.56	.86	.83	.76	.83	**.83**	.99	.97	.62	**.60**	**.66**	1.	1.	.69	.57	.62	1.	**.98**	.51	.70	.72	**.99**	.99	.88	.74	.67	**1.**	**1.**	.95	.87	.85
IResNet50 *CosFace*	*E*	9.4	1.5	0.3	0.2	**0.1**	22.0	13.9	7.3	5.5	4.9	39.5	**20.2**	**9.6**	8.6	7.7	34.9	17.2	5.9	**3.3**	**2.1**	27.9	8.5	**1.4**	0.9	0.6	10.3	**1.3**	0.4	**0.1**	**0.1**	27.3	10.2	3.5	**2.7**	2.8	8.8	3.0	1.6	1.4	1.3
	C	.98	.89	**.54**	.69	.76	.97	.99	.96	.92	.89	1.	1.	.91	**.82**	**.76**	1.	.99	.81	.64	**.54**	1.	1.	.95	.85	.76	1.	.97	.51	.69	.76	**1.**	**1.**	.97	**.89**	**.79**	.97	.92	.72	.62	.57
	κ	.65	.51	**.84**	**.89**	**.90**	.66	.89	.65	.56	.52	.86	.83	**.66**	**.74**	**.77**	.95	.90	.51	**.67**	**.72**	1.	1.	.83	.63	.54	1.	.97	.52	.66	.71	.99	.99	.94	**.85**	**.76**	.98	.95	.78	.70	.65
	L	1.	.97	.75	.61	.52	.99	**1.**	.99	.97	.96	1.	1.	.99	**.97**	.95	1.	.99	.91	.81	**.74**	1.	1.	.97	.92	.87	.99	.94	.52	.69	.77	**1.**	1.	.95	**.85**	**.76**	.94	.88	.64	.54	**.51**
	S	.69	.55	**.85**	**.90**	.89	.93	.91	.66	.56	.58	.60	.70	**.80**	**.85**	.83	.97	.96	.62	.60	.63	1.	.99	.62	**.62**	**.64**	1.	.96	.52	.71	.72	.98	.97	.87	.73	.67	.99	1.	**.97**	**.89**	**.87**
IResNet50 *MagFace*	*E*	8.6	**1.1**	**0.3**	0.3	0.1	21.3	13.4	**7.0**	5.3	**4.3**	39.6	20.6	9.8	8.5	**7.6**	**34.0**	**17.2**	**5.8**	3.5	2.3	27.4	**7.2**	1.5	**0.9**	0.8	10.4	1.5	0.4	0.2	**0.1**	27.5	**9.4**	**3.4**	2.8	2.7	8.6	**2.8**	**1.5**	**1.3**	1.3
	C	**.99**	.92	.53	**.70**	**.76**	**.99**	1.	.96	.92	**.89**	1.	1.	.91	.82	.75	**1.**	.99	.81	.63	.53	**1.**	1.	.96	.86	**.78**	**1.**	.99	.52	**.69**	**.76**	**1.**	**1.**	.97	.87	.78	**.99**	.95	.74	.64	**.59**
	κ	**.80**	.59	.82	.88	.90	.81	.93	.72	.62	**.57**	.96	.88	.59	.70	.74	.98	.93	.52	.65	.71	1.	1.	.85	.66	**.57**	1.	.99	.54	**.66**	**.72**	**1.**	.99	.94	.84	.76	**.99**	.96	.80	.71	**.66**
	L	**1.**	.98	.76	.61	.52	**1.**	1.	.99	.97	**.97**	**1.**	**1.**	.99	.97	**.96**	**1.**	1.	.91	.80	.73	1.	1.	.98	**.93**	**.88**	**1.**	.96	.50	**.69**	**.77**	**1.**	1.	.94	.84	.75	**.97**	.90	.65	.55	.50
	S	**.81**	**.62**	.84	.89	**.90**	.93	**.95**	**.75**	**.65**	.64	**.90**	**.87**	.69	.78	.79	**.99**	**.97**	**.68**	.53	.59	**1.**	**1.**	**.75**	.52	.54	**1.**	.97	**.55**	**.73**	**.74**	.99	**.99**	**.89**	**.76**	**.72**	1.	.99	.90	.81	.80
IResNet100 *ArcFace*	*E*	**8.5**	1.2	**0.3**	0.2	0.3	**20.2**	**12.8**	**6.1**	**5.1**	**4.2**	**38.0**	**19.3**	9.1	7.7	7.0	34.2	16.8	5.2	2.8	1.5	**26.2**	**6.7**	**1.2**	**0.7**	0.6	**9.6**	1.4	0.3	**0.1**	**0.1**	26.4	9.6	3.3	**2.6**	**2.4**	**8.1**	2.8	1.5	1.3	**1.2**
	C	.99	**.92**	**.54**	.65	**.75**	**.99**	.99	**.97**	**.94**	**.91**	**1.**	1.	**.93**	**.84**	.76	1.	**.99**	**.86**	**.69**	.55	**1.**	**1.**	**.97**	**.89**	.78	**1.**	**.99**	**.62**	.61	.73	**1.**	**1.**	**.99**	**.92**	.82	**.99**	**.95**	**.78**	**.68**	**.61**
	κ	**.83**	**.63**	.78	.87	.90	**.88**	**.96**	**.78**	**.68**	**.61**	**.98**	**.91**	.52	.67	.73	**.98**	**.95**	**.61**	.60	.69	**1.**	**1.**	**.91**	**.72**	**.58**	**1.**	**.99**	**.65**	.59	.69	1.	**1.**	**.96**	**.89**	**.79**	**.99**	**.97**	**.83**	**.75**	**.68**
	L	**1.**	**.98**	**.80**	**.67**	**.55**	1.	1.	**.99**	**.98**	**.97**	**1.**	1.	**.99**	**.98**	**.96**	1.	**1.**	**.93**	**.84**	.75	**1.**	**1.**	**.98**	**.94**	**.88**	**1.**	**.97**	**.59**	.61	.75	**1.**	**1.**	**.96**	**.89**	**.78**	**.97**	**.91**	**.70**	**.60**	**.52**
	S	.82	**.62**	.82	.89	.88	.94	.94	.72	.61	.58	.88	.84	.72	.81	**.82**	**.99**	.97	.66	**.57**	**.64**	1.	1.	.71	.55	.63	1.	**.98**	**.57**	.66	.70	**.99**	.99	.90	.78	.70	**1.**	**1.**	.97	.91	.87
IResNet100 *CosFace*	*E*	8.7	1.1	0.3	0.3	0.1	21.8	13.5	7.2	5.2	4.2	38.7	19.6	9.1	**7.5**	7.0	34.2	16.4	**5.0**	**2.5**	1.6	28.6	7.7	1.4	0.8	**0.5**	10.9	1.4	**0.2**	0.2	**0.1**	26.7	9.7	3.4	2.7	2.5	8.5	2.9	1.6	1.3	1.3
	C	.98	.89	.51	**.66**	.75	.97	.99	.97	.93	.90	1.	1.	.92	.84	**.76**	1.	.99	.83	.68	**.56**	**1.**	1.	.96	.88	.77	1.	.98	.56	**.64**	.74	**1.**	**1.**	.98	.92	**.82**	.97	.92	.74	.66	.59
	κ	.69	.53	**.82**	**.88**	**.90**	.71	.90	.71	.62	.57	.90	.84	**.62**	**.72**	**.76**	.96	.91	.54	**.63**	**.71**	1.	1.	.87	.67	.55	1.	.98	.58	**.61**	.70	.99	.99	.95	.88	.79	.98	.95	.80	.73	.67
	L	1.	.97	.77	.65	.54	1.	1.	.99	.98	.97	1.	1.	.99	.97	.96	1.	.99	.92	.84	**.75**	1.	1.	.97	.94	.88	.99	.94	.54	**.64**	**.75**	1.	1.	.96	.88	.78	.94	.88	.66	.58	.51
	S	.72	.56	**.84**	**.89**	.88	**.95**	.92	.72	.60	.60	.66	.72	**.77**	**.83**	.82	.98	.96	.66	.56	.61	1.	.99	.62	**.61**	**.65**	1.	.95	.51	**.69**	.71	.98	.97	.88	.77	.69	.99	1.	**.98**	**.92**	**.90**
IResNet100 *MagFace*	*E*	8.8	**1.1**	**0.3**	**0.1**	**0.1**	22.0	12.9	6.4	**5.1**	4.3	40.0	20.3	**8.9**	7.9	**7.0**	**33.8**	**16.3**	5.0	3.0	**1.4**	27.9	7.0	1.2	**0.7**	0.7	10.1	**1.3**	0.3	0.2	0.1	**26.2**	**9.0**	**3.0**	2.7	**2.4**	8.5	**2.7**	**1.4**	**1.3**	1.2
	C	**.99**	.91	.51	.65	.75	.99	**1.**	.97	.94	.91	1.	**1.**	.92	.84	.76	**1.**	.99	.83	.68	.55	**1.**	1.	.96	.88	**.78**	**1.**	.99	.57	.63	**.74**	**1.**	1.	.98	.92	.80	.99	.94	.76	.68	.61
	κ	.81	.58	.81	.87	.90	.86	.94	.76	.67	.60	.96	.89	.55	.67	.73	.98	.93	.57	.61	.70	**1.**	1.	.89	.71	.58	1.	.98	.59	.61	**.70**	**1.**	.99	.95	.88	.78	.99	.96	.82	.74	.68
	L	1.	.98	.78	.66	.54	**1.**	**1.**	.99	.98	.97	1.	**1.**	.99	.98	.96	**1.**	1.	.92	.83	.74	1.	1.	.98	.94	.88	1.	.96	.54	.63	.75	**1.**	1.	.96	.88	.78	.97	.89	.67	.59	.52
	S	**.83**	.61	.82	.88	**.89**	.94	**.95**	**.78**	**.69**	**.65**	**.91**	**.87**	.67	.77	.79	.99	**.97**	**.71**	.51	.59	**1.**	**1.**	**.76**	.52	.56	**1.**	.97	.51	.68	**.71**	.99	**.99**	**.90**	**.81**	**.74**	.99	.99	.93	.86	.83
ViT-T *CosFace*	*E*	9.7	1.8	0.3	0.3	0.3	19.2	10.0	5.2	5.2	4.2	38.1	24.7	11.8	9.8	8.8	34.1	19.8	6.6	4.5	3.1	28.9	13.9	2.7	1.9	1.3	11.7	2.3	0.4	0.3	0.2	29.2	13.3	4.3	3.6	3.0	8.7	3.9	1.7	1.6	1.4
	C	.98	.80	.65	.70	.76	.98	.98	.90	.88	.85	1.	.99	.88	.84	.77	1.	.97	.72	.65	.55	1.	1.	.95	.90	.82	1.	.92	.57	.63	.71	1.	1.	.95	.91	.83	.96	.88	.65	.62	.57
	κ	.62	.65	.88	.89	.90	.67	.67	.53	.52	.51	.78	.65	.68	.71	.74	.94	.79	.60	.65	.70	1.	1.	.80	.72	.61	1.	.92	.55	.61	.67	.99	.98	.90	.87	.79	.97	.92	.72	.69	.65
	L	1.	.93	.66	.60	.52	1.	.99	.97	.96	.95	1.	1.	.98	.97	.96	1.	.99	.86	.82	.75	1.	1.	.97	.95	.91	1.	.87	.58	.64	.73	1.	1.	.92	.87	.80	.93	.81	.57	.54	.51
	S	.67	.58	.88	.89	.89	.85	.76	.57	.54	.54	.62	.51	.79	.82	.80	.97	.93	.58	.54	.60	.99	.99	.65	.50	.56	.99	.96	.57	.67	.69	.98	.96	.84	.76	.70	.99	1.	.92	.87	.86

though it remains above chance even on the held-out WebFace4M split (Table 5, discussed in Section 5.4).

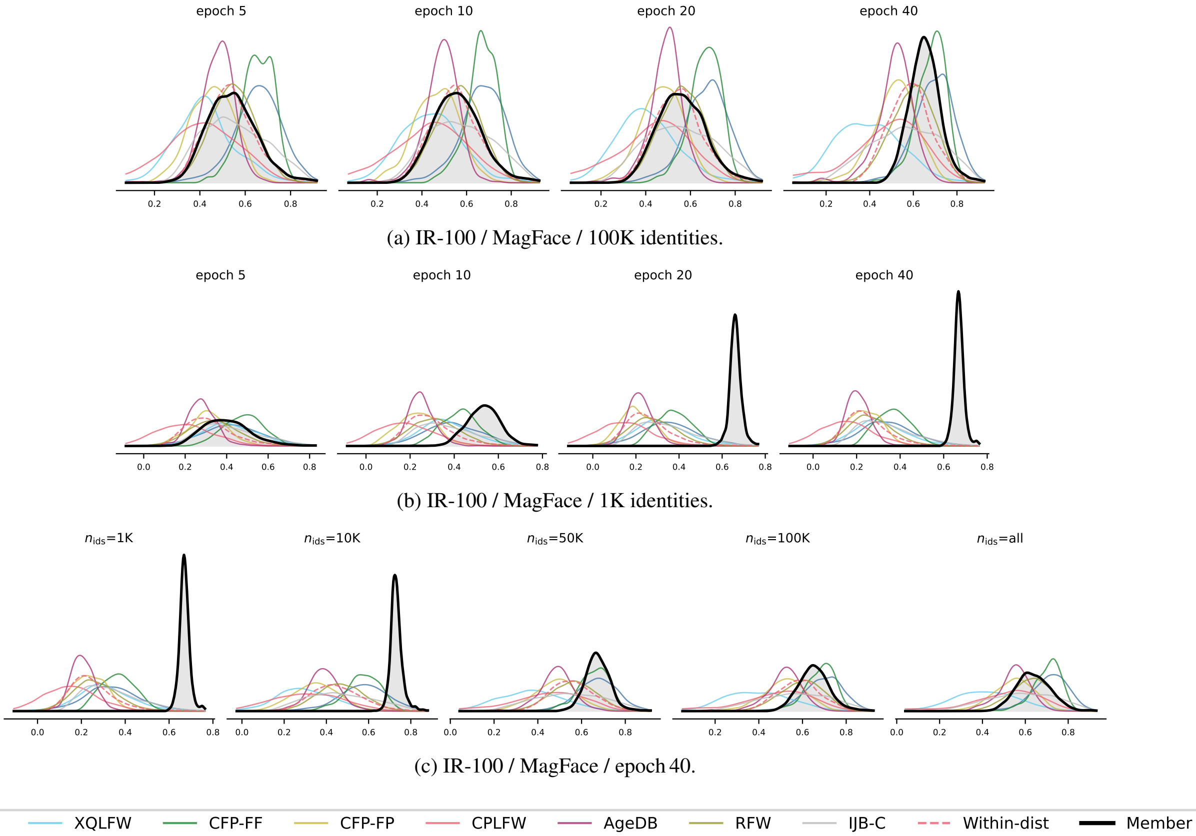
Panels (a) and (b) of Figure 2 isolate the effect of training duration at two extremes of identity coverage. At n_{ids}=1K, the PairCos separation between members and non-members increases with the number of training epochs. At n_{ids}=100K the same trend is present but attenuated.

Table 3 reports the ANOVA results ($\hat{\eta}_p^2$, the proportion of variance explained after accounting for all other factors). Training-set size explains the most variance ($\hat{\eta}_p^2 \geq .95$), followed by training duration ($\hat{\eta}_p^2 \geq .89$). The strong $n_{\text{ids}} \times$ Epoch interaction ($\hat{\eta}_p^2 \geq .92$) further shows that the epoch effect is moderated by identity coverage. The $n_{\text{ids}} \times$ Epoch cell means on the same-domain WebFace4M split show that the mean PairCos AUC increase from epoch 5 to epoch 40 ranges from +.38 at n_{ids}=10K down to +.16 at n_{ids}=100K, with n_{ids}=1K reaching a ceiling near 1.0 already at epoch 10. Models with small identity pools thus show the largest separability gains over training, whereas models trained on 100K identities have only a modest increase.

Backbone and loss head. The ANOVA also shows that backbone architecture and loss head explain far less variance in threshold-classifier AUC, the loss-head factor reaching at most $\hat{\eta}_p^2 = .04$ and backbone depth at most .09 across all four statistics — an order of magnitude below the contributions of n_{ids} and number of epochs. Beyond the IResNet grid, a ViT-T [11, 32] backbone (CosFace) trained under the identical protocol [2] reproduces the same n_{ids} trend in both EER and membership AUC (Table 4, bottom block), so the effect is not specific to convolutional backbones.

5.3. Effect of Input-Data Distribution

The benchmarks used in this study differ in pose, quality, age, and demographics, which affect embedding-cluster geometry independently of membership. The distributional-shift rows of Table 3 quantify these effects, and we discuss each factor below. For matched-pair benchmarks, we also



(a) IR-100 / MagFace / 100K identities.

(b) IR-100 / MagFace / 1K identities.

(c) IR-100 / MagFace / epoch 40.

Figure 2. PAIRCOS ($\bar{s}$) kernel density estimates. The bold black curve (with shading) shows *member* identities (present in the training set). All other curves show *non-member* identities: the eight thin solid lines correspond to cross-distribution benchmarks whose domain differs from the training data (one colour per dataset; see legend), while the dashed line ("Within-dist") shows non-members from the held-out WebFace4M split, which shares the training domain (no domain shift). (a, b) vary training epochs at fixed n_{ids}; (c) varies n_{ids} at epoch 40.

translate each effect into the AUC change it would produce at the same-domain training-attributable baseline at n_{ids}=100K (where the AUC is ≈ 0.71). We do so by passing each AUC through the inverse standard normal cumulative distribution function (a probit transform), on which equal shifts in the underlying non-member distribution correspond to equal numerical changes regardless of baseline. We take the matched-pair difference on this transformed scale and convert it back to AUC at the chosen baseline. This step is needed because the AUC scale saturates near 1, so a fixed shift in the underlying distribution produces a smaller AUC change near the ceiling than in the middle.

Pose. Profile-view faces produce large shifts in cluster-statistic distributions. The CFP-FP vs. CFP-FF comparison is the cleanest controlled experiment: both sets share the same 500 identities, so the observed Cohen's d' (standardised mean difference) is attributable to pose alone (Table 3). CPLFW vs. LFW yields even larger $|d'|$ for PAIRCOS and PENLOGIT, but those two sets are not identity-matched, so other factors may contribute. In both comparisons, profile views widen the intra-identity distribution regardless of membership. This pushes statistic values towards the non-member range and inflates separability. At the n_{ids}=100K baseline, the pose contribution from CFP-FF to CFP-FP corresponds to an AUC change of $+0.26$, matching or exceeding the training contribution above chance ($+0.21$).

Image quality. XQLFW shares LFW's identity set and pair protocol but degrades one image per pair. For PAIRCOS and PENLOGIT, the resulting Cohen's d' is of the same order as that of cross-pose variation though somewhat smaller; κ shows a smaller quality effect ($|d'|$=0.48 vs. 1.13 for pose). Lower image quality increases intra-identity spread in embedding space and shifts cluster-statistic distributions in the same direction as pose variation. At the same baseline, the quality contribution from LFW to XQLFW corresponds to an AUC change of $+0.22$, comparable to the pose contribution.

Age. Within AgeDB-30, the Spearman rank correlation ρ (a non-parametric measure of monotonic association) between within-identity age spread and each cluster statistic (Table 3) is weak and varies in sign across statistics. Only κ exhibits the expected negative association ($\rho = -0.26$), under which identities spanning wider age ranges form less concentrated clusters, whereas PAIRCOS is essentially un-

Table 5. **Statistic fusion recovers more membership information than any single statistic.** Each cell shows best single-statistic AUC / fold-averaged MLP AUC. **Bold** highlights the MLP gain over the best single statistic.

		n_{ids}			
Backbone	Loss	1K	10K	50K	100K
IR-34	Arc	1.00 / **1.00**	.98 / **1.00**	.80 / **.87**	.65 / **.69**
	Cos	.99 / **1.00**	.98 / **.99**	.79 / **.88**	.65 / **.71**
	Mag	1.00 / **1.00**	.98 / **1.00**	.79 / **.85**	.64 / **.68**
IR-50	Arc	1.00 / **1.00**	.99 / **1.00**	.84 / **.91**	.68 / **.73**
	Cos	.99 / **1.00**	.98 / **.99**	.82 / **.91**	.68 / **.75**
	Mag	1.00 / **1.00**	.98 / **1.00**	.82 / **.89**	.67 / **.72**
IR-100	Arc	1.00 / **1.00**	.99 / **1.00**	.86 / **.92**	.73 / **.79**
	Cos	.99 / **1.00**	.98 / **.99**	.84 / **.92**	.71 / **.80**
	Mag	1.00 / **1.00**	.98 / **1.00**	.85 / **.91**	.72 / **.79**

correlated ($\rho = +0.01$) and both PENLOGIT and PROTOCE carry small positive correlations ($\rho = +0.17$ and $\rho = +0.33$). Age-range variation therefore exerts the weakest and least consistent influence on cluster geometry among the confounds examined here.

Ethnicity. A Kruskal–Wallis test (non-parametric one-way ANOVA) across the four RFW demographic groups (African, Asian, Caucasian, Indian) yields $\epsilon^2 \approx 0.02$ (small) for PAIRCOS, κ, and PENLOGIT, and $\epsilon^2 = 0.12$ (medium) for PROTOCE. Cluster-statistic values therefore vary across demographic groups within the non-member pool, with PROTOCE again being the most sensitive statistic. Although these effect sizes are modest relative to the n_{ids} factor, they indicate that the demographic composition of a non-member benchmark affects the baseline against which membership is measured.

Pose and quality produce the largest shifts in cluster-statistic distributions, with demographic composition and age contributing smaller and, for age, less consistent effects. The matched-pair pose and quality contributions show that the domain-induced component can match or exceed the training contribution itself at higher n_{ids}, so the threshold-classifier AUC on any non-WebFace4M benchmark cannot be interpreted as membership signal alone.

5.4. Statistic Fusion

To test whether combining statistics recovers more membership information, we train a two-hidden-layer MLP (64→32 units, ReLU, BatchNorm, dropout 0.3; Adam with $\eta = 10^{-3}$) on all four cluster statistics. For each epoch-40 configuration we perform five-fold stratified cross-validation with an 80/20 identity split on same-domain WebFace4M identities, reserving 15% of each training fold as a validation set for early stopping. Table 5 reports the best single-statistic AUC alongside the fold-averaged MLP AUC. The MLP consistently matches or outperforms the best individual statistic, with the largest gains at n_{ids}=50K and 100K where individual statistics are weakest (one-sided Wilcoxon signed-rank: W=666, $p < 10^{-10}$).

5.5. Limitations

Generalisation across training data. All experiments use a single training source (WebFace4M). Whether the trends generalise to other training datasets is an open question, as comparably large datasets are not publicly available or have been retracted.

Experimental controls. We do not evaluate defences such as DP-SGD [1] or knowledge distillation [17]. The number of probe images k per identity varies across benchmarks and is not controlled for, which may affect estimates of the cluster statistics non-uniformly. The held-out split is constructed by random partitioning; in practice, membership boundaries may correlate with identity-level attributes.

Statistical assumptions and scope. Because our four statistics capture only cluster geometry, the measured separability is a lower bound on the membership information available from embeddings; a richer model operating on raw embedding vectors could extract more. The probit-scale translation in Section 5.3 assumes approximately Gaussian member and non-member statistic distributions and stable variances across matched pairs.

6. Conclusion and Future Work

We evaluated how much membership information is retained in the embedding geometry of 180 open-set face recognition models across nine benchmarks. The results lead to three practical conclusions. First, training-set size has the largest effect on member/non-member separability, while backbone architecture and loss head contribute far less. Training on more identities is therefore the only design choice we tested that substantially reduces the geometric membership signal. Second, cross-domain non-member benchmarks inflate the apparent signal by conflating domain shift with training-attributable differences. Privacy audits should therefore use same-domain held-out references. Third, even at high n_{ids} where individual statistics are weak, fusing all statistics with a learned classifier recovers additional membership information (Table 5). As our statistics capture only cluster geometry, the measured separability is a lower bound on the membership information available from embeddings. Future work includes extending these benchmarks to other backbones and evaluating whether defences such as [1] can reduce the geometric signal without degrading performance.

Acknowledgments

This work has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101189650 (CERTAIN: *Certification for Ethical and Regulatory Transparency in Artificial Intelligence*), and the Swiss State Secretariat for Education, Research and Innovation (SERI).

References

[1] M. Abadi, A. Chu, I. J. Goodfellow, H. B. McMahan, I. Mironov, K. Talwar, and L. Zhang. Deep learning with differential privacy. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, CCS 2016, Vienna, Austria, October 24-28, 2016*, pages 308–318. ACM, 2016. 3, 8

[2] X. An, J. Deng, J. Guo, Z. Feng, X. Zhu, J. Yang, and T. Liu. Killing two birds with one stone: Efficient and robust training of face recognition cnns by partial FC. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2022, New Orleans, LA, USA, June 18-24, 2022*, pages 4032–4041. IEEE, 2022. 6

[3] A. Banerjee, I. S. Dhillon, J. Ghosh, and S. Sra. Clustering on the unit hypersphere using von mises-fisher distributions. *J. Mach. Learn. Res.*, 6:1345–1382, 2005. 3

[4] L. Bourtoule, V. Chandrasekaran, C. A. Choquette-Choo, H. Jia, A. Travers, B. Zhang, D. Lie, and N. Papernot. Machine unlearning. In *42nd IEEE Symposium on Security and Privacy, SP 2021, San Francisco, CA, USA, May 23-27, 2021*, pages 141–159. IEEE, 2021. 3

[5] N. Carlini, S. Chien, M. Nasr, S. Song, A. Terzis, and F. Tramèr. Membership inference attacks from first principles. In *43rd IEEE Symposium on Security and Privacy, SP 2022, San Francisco, CA, USA, May 22-26, 2022*, pages 1897–1914. IEEE, 2022. 2

[6] G. Chen, Y. Zhang, and F. Song. SLMIA-SR: Speaker-level membership inference attacks against speaker recognition systems. In *31st Annual Network and Distributed System Security Symposium, NDSS 2024, San Diego, CA, USA, February 26 - March 1, 2024*. The Internet Society, 2024. 2

[7] M. Chen, Z. Zhang, T. Wang, M. Backes, and Y. Zhang. FACE-AUDITOR: Data auditing in facial recognition systems. In *32nd USENIX Security Symposium, USENIX Security 2023, Anaheim, CA, USA, August 9-11, 2023*, pages 7195–7212. USENIX Association, 2023. 2

[8] D. DeAlcala, G. Mancera, A. Morales, J. Fiérrez, R. Tolosana, and J. Ortega-Garcia. A comprehensive analysis of factors impacting membership inference. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, CVPRW 2024, Seattle, WA, USA, June 17-18, 2024*, pages 3585–3593. IEEE, 2024. 2

[9] D. DeAlcala, A. Morales, J. Fierrez, G. Mancera, R. Tolosana, and J. Ortega-Garcia. Is my data in your AI? Membership inference test (MINT) applied to face biometrics. *IEEE Access*, 13:163805–163819, 2025. 2

[10] J. Deng, J. Guo, N. Xue, and S. Zafeiriou. Arcface: Additive angular margin loss for deep face recognition. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019, Long Beach, CA, USA, June 16-20, 2019*, pages 4690–4699. Computer Vision Foundation / IEEE, 2019. 1, 3, 4

[11] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby. An image is worth 16x16 words: Transformers for image recognition at scale. In *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021*. OpenReview.net, 2021. 6

[12] V. Feldman. Does learning require memorization? A short tale about a long tail. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing, STOC 2020, Chicago, IL, USA, June 22-26, 2020*, pages 954–959. ACM, 2020. 3

[13] M. Fredrikson, S. Jha, and T. Ristenpart. Model inversion attacks that exploit confidence information and basic countermeasures. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security, CCS 2015, Denver, CO, USA, October 12-16, 2015*, pages 1322–1333. ACM, 2015. 3

[14] Regulation (EU) 2016/679 of the European Parliament and of the Council (general data protection regulation). Official Journal of the European Union, L 119, 1–88, 2016. Art. 17 (Right to erasure). 2

[15] A. Golatkar, A. Achille, and S. Soatto. Eternal sunshine of the spotless net: Selective forgetting in deep networks. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2020, Seattle, WA, USA, June 13-19, 2020*, pages 9301–9309. Computer Vision Foundation / IEEE, 2020. 3

[16] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, June 27-30, 2016*, pages 770–778. IEEE Computer Society, 2016. 4

[17] G. E. Hinton, O. Vinyals, and J. Dean. Distilling the knowledge in a neural network. *CoRR*, abs/1503.02531, 2015. NeurIPS Deep Learning Workshop. 8

[18] H. Hu, Z. Salcic, L. Sun, G. Dobbie, P. S. Yu, and X. Zhang. Membership inference attacks on machine learning: A survey. *ACM Comput. Surv.*, 54(11s):1–37, 2022. 2

[19] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical Report 07-49, University of Massachusetts, Amherst, 2007. 4, 5

[20] Y. Huang, H. Chen, Y. Wang, and L. Wang. Inference attacks against face recognition model without classification layers. *CoRR*, abs/2401.13719, 2024. 2, 3

[21] M. Knoche, S. Hörmann, and G. Rigoll. Cross-quality LFW: A database for analyzing cross- resolution image face recognition in unconstrained environments. In *16th IEEE International Conference on Automatic Face and Gesture Recognition, FG 2021, Jodhpur, India, December 15-18, 2021*, pages 1–5. IEEE, 2021. 4, 5

[22] G. Li, S. Rezaei, and X. Liu. User-level membership inference attack against metric embedding learning. *CoRR*, abs/2203.02077, 2022. 2, 3

[23] H. Liu, J. Jia, W. Qu, and N. Z. Gong. EncoderMI: Membership inference against pre-trained encoders in contrastive learning. In *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security, CCS 2021, Virtual Event, Republic of Korea, November 15-19, 2021*, pages 2081–2095. ACM, 2021. 2

[24] G. Mancera, D. DeAlcala, A. Morales, R. Tolosana, and J. Fierrez. Membership inference test: Auditing training data in object classification models. In *Deployable AI Workshop (DAI 2025), co-located with the 39th AAAI Conference on Artificial Intelligence (AAAI-25)*, 2025. 2

[25] B. Maze, J. C. Adams, J. A. Duncan, N. D. Kalka, T. Miller, C. Otto, A. K. Jain, W. T. Niggel, J. Anderson, J. Cheney, and P. Grother. IARPA janus benchmark - C: face dataset and protocol. In *2018 International Conference on Biometrics, ICB 2018, Gold Coast, Australia, February 20-23, 2018*, pages 158–165. IEEE, 2018. 4, 5

[26] Q. Meng, S. Zhao, Z. Huang, and F. Zhou. Magface: A universal representation for face recognition and quality assessment. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2021, virtual, June 19-25, 2021*, pages 14225–14234. Computer Vision Foundation / IEEE, 2021. 1, 3, 4

[27] S. Moschoglou, A. Papaioannou, C. Sagonas, J. Deng, I. Kotsia, and S. Zafeiriou. Agedb: The first manually collected, in-the-wild age database. In *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops, CVPR Workshops 2017, Honolulu, HI, USA, July 21-26, 2017*, pages 1997–2005. IEEE Computer Society, 2017. 4, 5

[28] S. Rezaei and X. Liu. On the difficulty of membership inference attacks. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2021, virtual, June 19-25, 2021*, pages 7892–7900. Computer Vision Foundation / IEEE, 2021. 3

[29] A. Salem, Y. Zhang, M. Humbert, P. Berrang, M. Fritz, and M. Backes. ML-Leaks: Model and data independent membership inference attacks and defenses on machine learning models. In *26th Annual Network and Distributed System Security Symposium, NDSS 2019, San Diego, CA, USA, February 24-27, 2019*. The Internet Society, 2019. 2

[30] S. Sengupta, J. Chen, C. D. Castillo, V. M. Patel, R. Chellappa, and D. W. Jacobs. Frontal to profile face verification in the wild. In *2016 IEEE Winter Conference on Applications of Computer Vision, WACV 2016, Lake Placid, NY, USA, March 7-10, 2016*, pages 1–9. IEEE Computer Society, 2016. 4, 5

[31] R. Shokri, M. Stronati, C. Song, and V. Shmatikov. Membership inference attacks against machine learning models. In *2017 IEEE Symposium on Security and Privacy, SP 2017, San Jose, CA, USA, May 22-26, 2017*, pages 3–18. IEEE Computer Society, 2017. 2

[32] H. Touvron, M. Cord, M. Douze, F. Massa, A. Sablayrolles, and H. Jégou. Training data-efficient image transformers & distillation through attention. In M. Meila and T. Zhang, editors, *Proceedings of the 38th International Conference on Machine Learning, ICML 2021, 18-24 July 2021, Virtual Event*, volume 139 of *Proceedings of Machine Learning Research*, pages 10347–10357. PMLR, 2021. 6

[33] H. Wang, Y. Wang, Z. Zhou, X. Ji, D. Gong, J. Zhou, Z. Li, and W. Liu. Cosface: Large margin cosine loss for deep face recognition. In *2018 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2018, Salt Lake City, UT, USA, June 18-22, 2018*, pages 5265–5274. Computer Vision Foundation / IEEE Computer Society, 2018. 1, 3, 4

[34] M. Wang, W. Deng, J. Hu, X. Tao, and Y. Huang. Racial faces in the wild: Reducing racial bias by information maximization adaptation network. In *2019 IEEE/CVF International Conference on Computer Vision, ICCV 2019, Seoul, Korea (South), October 27 - November 2, 2019*, pages 692–702. IEEE, 2019. 4, 5

[35] T. Zheng and W. Deng. Cross-pose LFW: A database for studying cross-pose face recognition in unconstrained environments. Technical Report 18-01, Beijing University of Posts and Telecommunications, February 2018. 4, 5

[36] Z. Zhu, G. Huang, J. Deng, Y. Ye, J. Huang, X. Chen, J. Zhu, T. Yang, D. Du, J. Lu, and J. Zhou. Webface260m: A benchmark for million-scale deep face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.*, 45(2):2627–2644, 2023. 1, 4

Supplementary Material:
Quantifying Training Membership Information in the Hyperspherical Embedding Geometry of Face Recognition Models

Ünsal Öztürk[1] Sébastien Marcel[1,2]

[1]Idiap Research Institute, Martigny, Switzerland
[2]Université de Lausanne (UNIL), Lausanne, Switzerland

{unsal.ozturk, sebastien.marcel}@idiap.ch

Contents

Abstract

This document provides extended results that complement the main paper. It is organized as follows. Sections 1 and 2 present full ROC and DET curves for, respectively, face-verification performance and the threshold-classifier membership-inference attack, covering all 180 model configurations (9 backbone×loss combinations × 5 training-set sizes × 4 training epochs). Curves are grouped in three views each: by evaluation dataset, by backbone and loss head, and by training-set size. Section 3 gives per-epoch tables of verification EER and threshold-classifier AUC for all four per-identity geometric statistics (PairCos, vMF-κ, PenLogit, ProtoCE) across every configuration and benchmark. Sections 4–7 provide kernel-density score-distribution atlases for each of the four statistics, with subsections varying backbone, loss head, training-set size, and evaluation dataset in turn.

1. Verification ROC and DET curves

Each figure pairs a ROC curve (left) and a DET curve (right) for a given grouping of the 180 model configurations. Subsection 1.1 groups by evaluation dataset; Subsection 1.2 groups by backbone and loss head; Subsection 1.3 groups curves by training-set size (n_{ids}) on each dataset.

1.1. Grouped by evaluation dataset

AgeDB-30

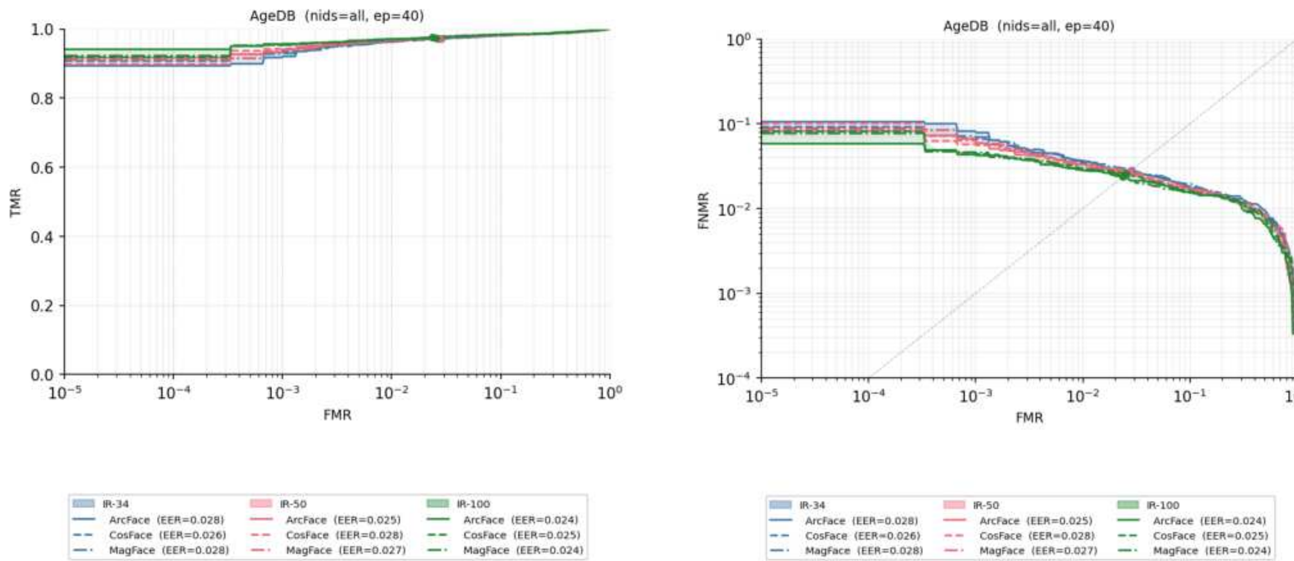


Figure 1. **ROC and DET curves — AgeDB-30.** Each curve represents one model configuration (backbone $\times$ loss head $\times$ n_{ids} $\times$ epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

CFP-FF

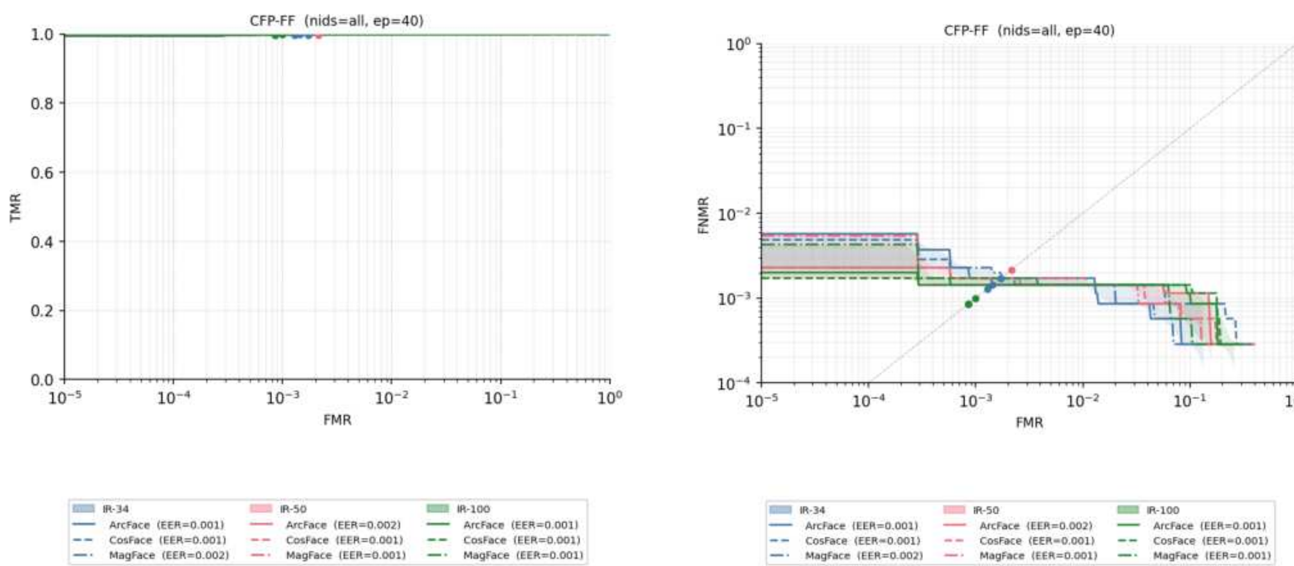


Figure 2. **ROC and DET curves — CFP-FF.** Each curve represents one model configuration (backbone $\times$ loss head $\times$ n_{ids} $\times$ epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

CFP-FP

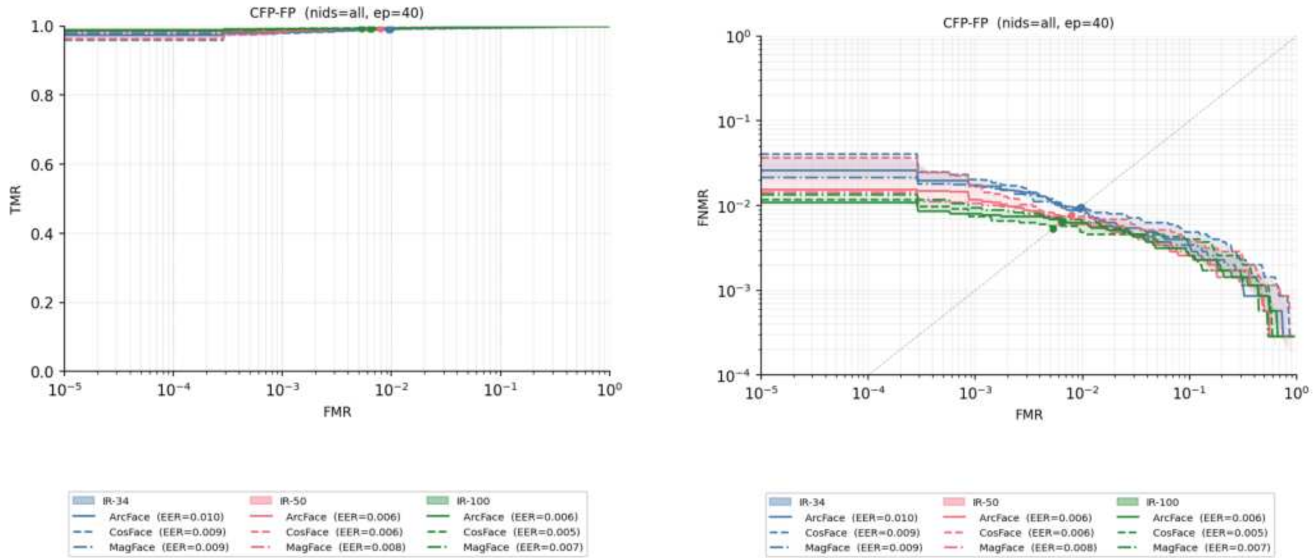


Figure 3. **ROC and DET curves — CFP-FP.** Each curve represents one model configuration (backbone $\times$ loss head $\times$ n_{ids} $\times$ epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

CPLFW

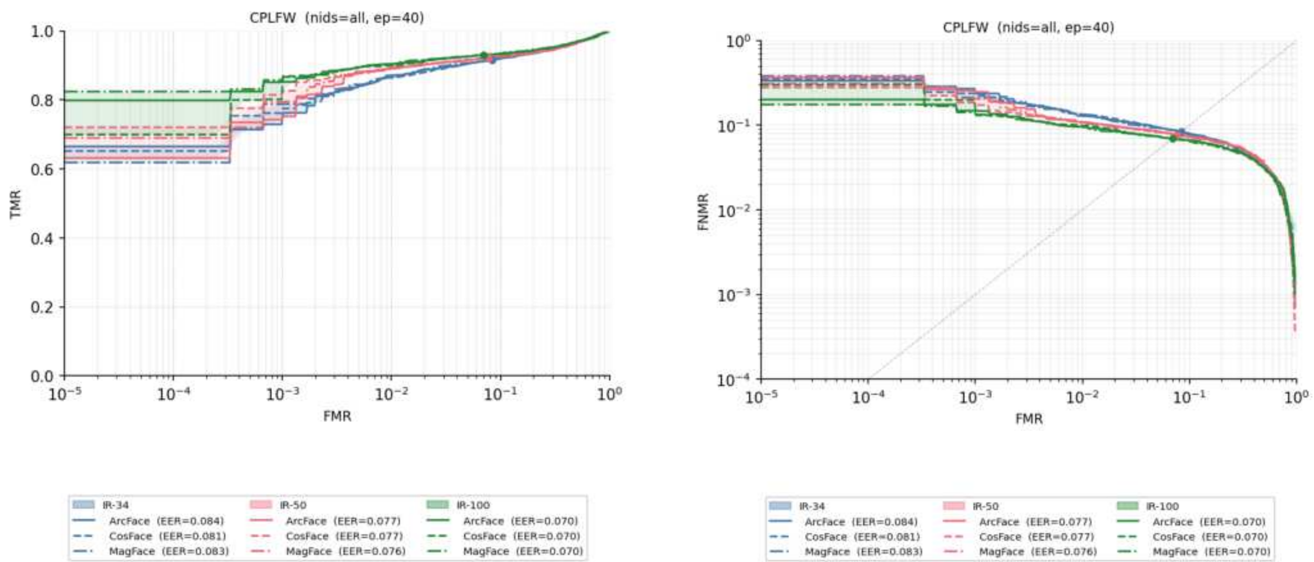


Figure 4. **ROC and DET curves — CPLFW.** Each curve represents one model configuration (backbone $\times$ loss head $\times$ n_{ids} $\times$ epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IJB-C

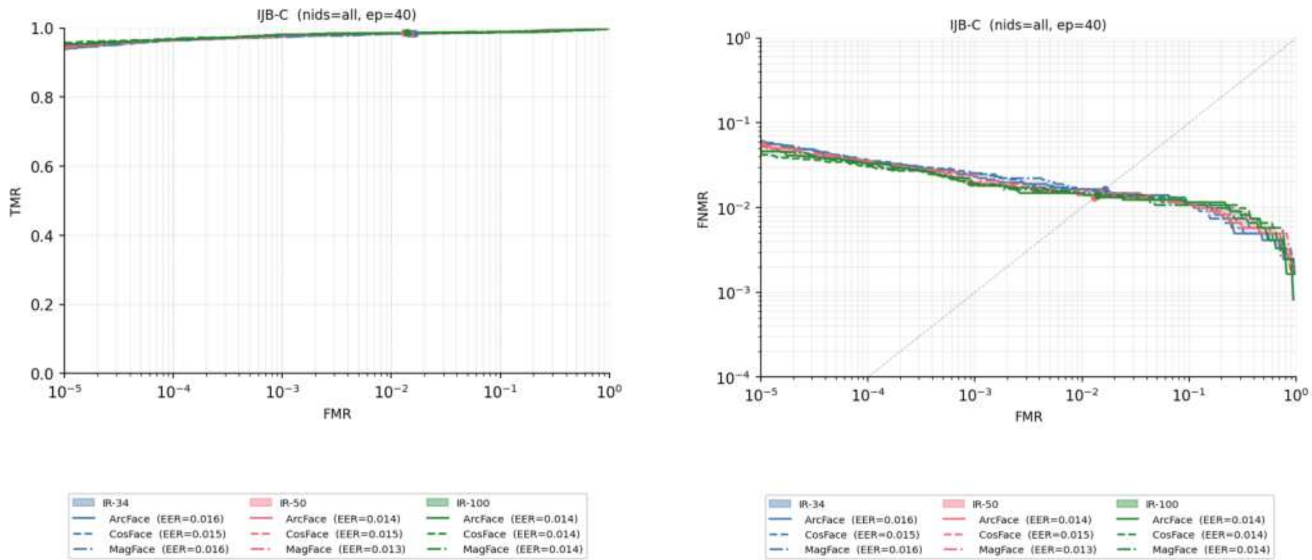


Figure 5. **ROC and DET curves — IJB-C.** Each curve represents one model configuration (backbone $\times$ loss head $\times$ n_{ids} $\times$ epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

LFW

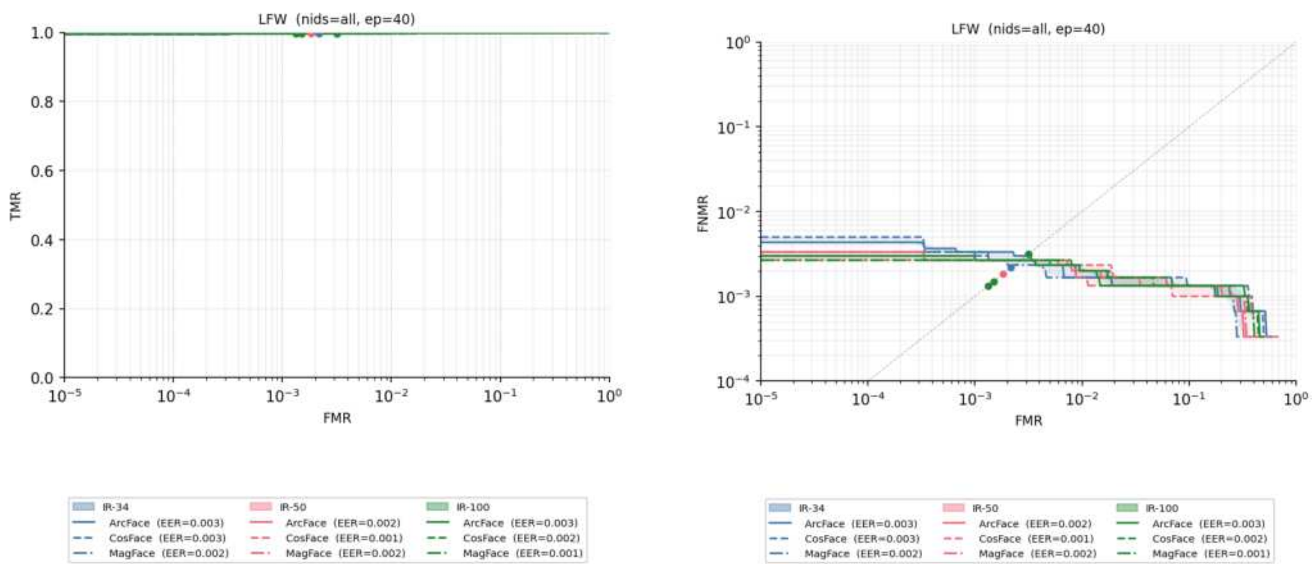


Figure 6. **ROC and DET curves — LFW.** Each curve represents one model configuration (backbone $\times$ loss head $\times$ n_{ids} $\times$ epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

RFW

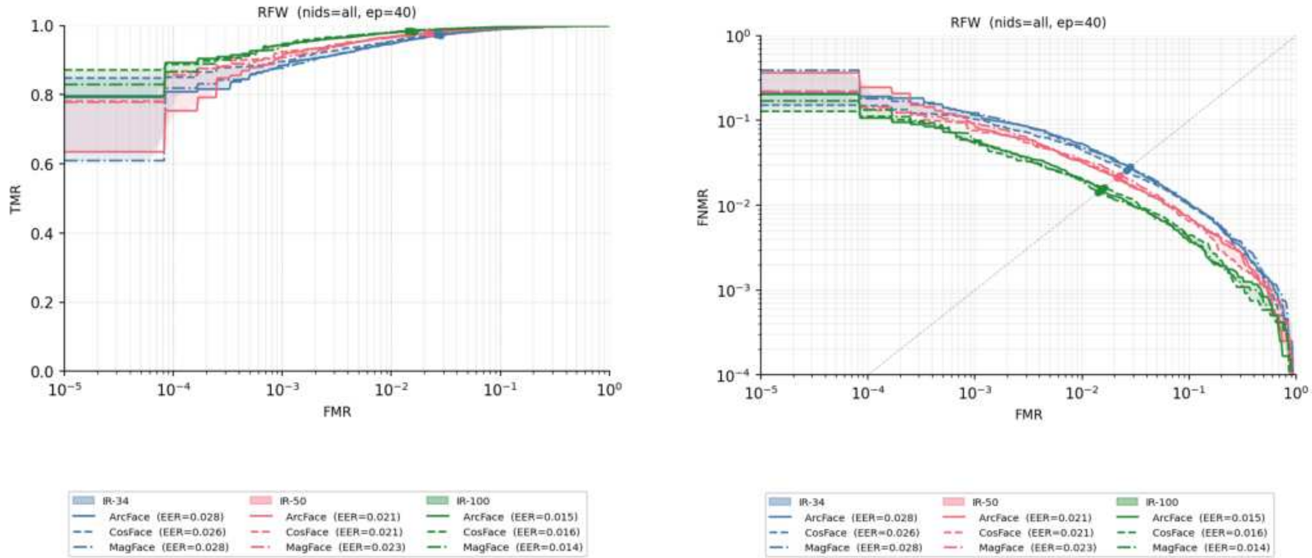


Figure 7. **ROC and DET curves — RFW.** Each curve represents one model configuration (backbone × loss head × n_{ids} × epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

XQLFW

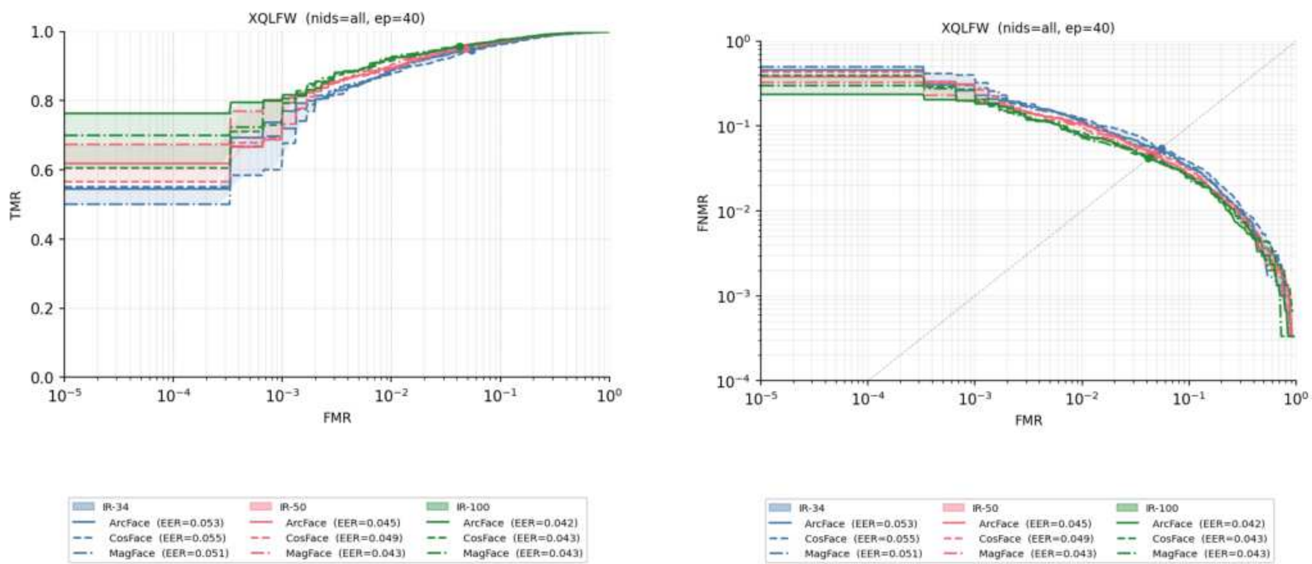


Figure 8. **ROC and DET curves — XQLFW.** Each curve represents one model configuration (backbone × loss head × n_{ids} × epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

1.2. Grouped by backbone and loss head

IResNet-34 / ArcFace

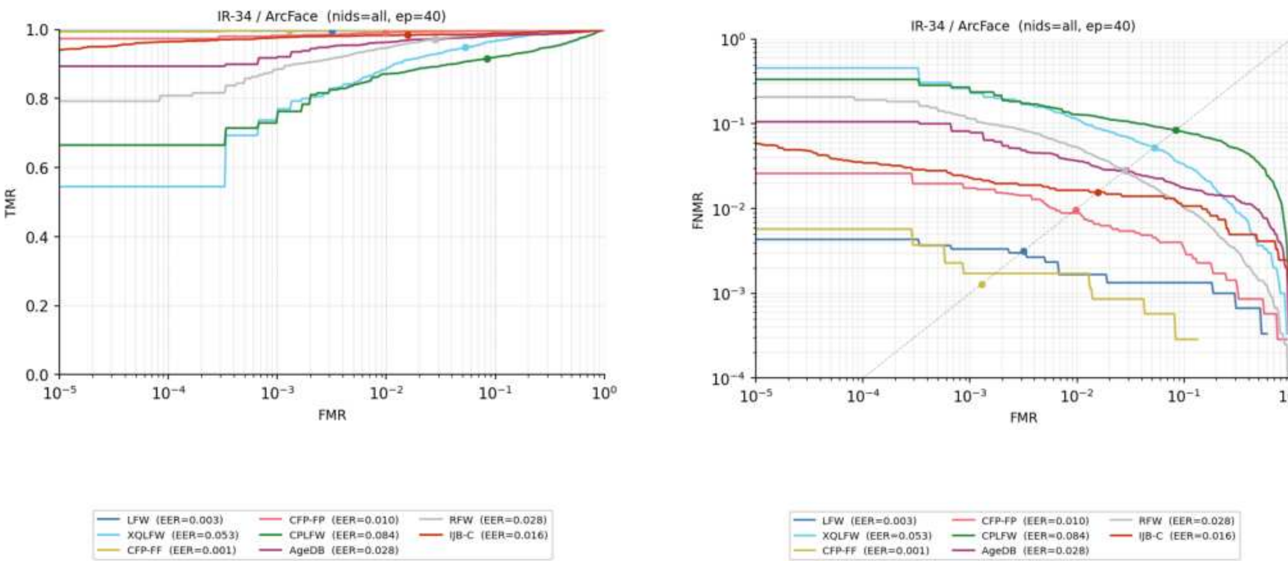


Figure 9. **ROC and DET curves — IResNet-34 / ArcFace.** Each curve represents one combination of training-set size (n_{ids}) and training epoch across all evaluation benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-34 / CosFace

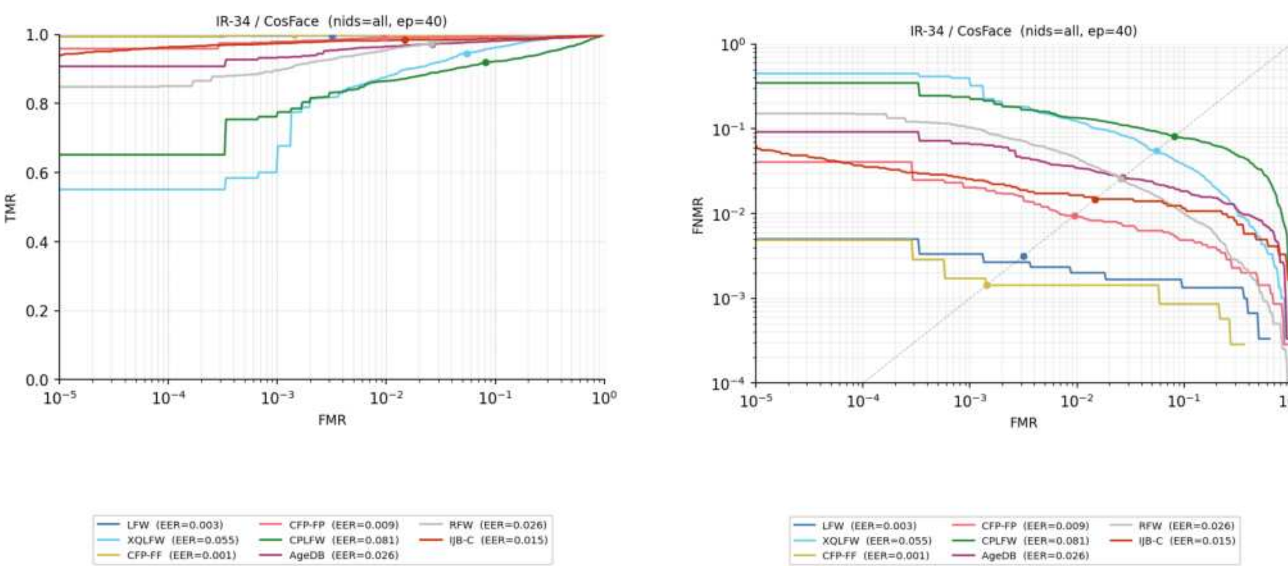


Figure 10. **ROC and DET curves — IResNet-34 / CosFace.** Each curve represents one combination of training-set size (n_{ids}) and training epoch across all evaluation benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-34 / MagFace

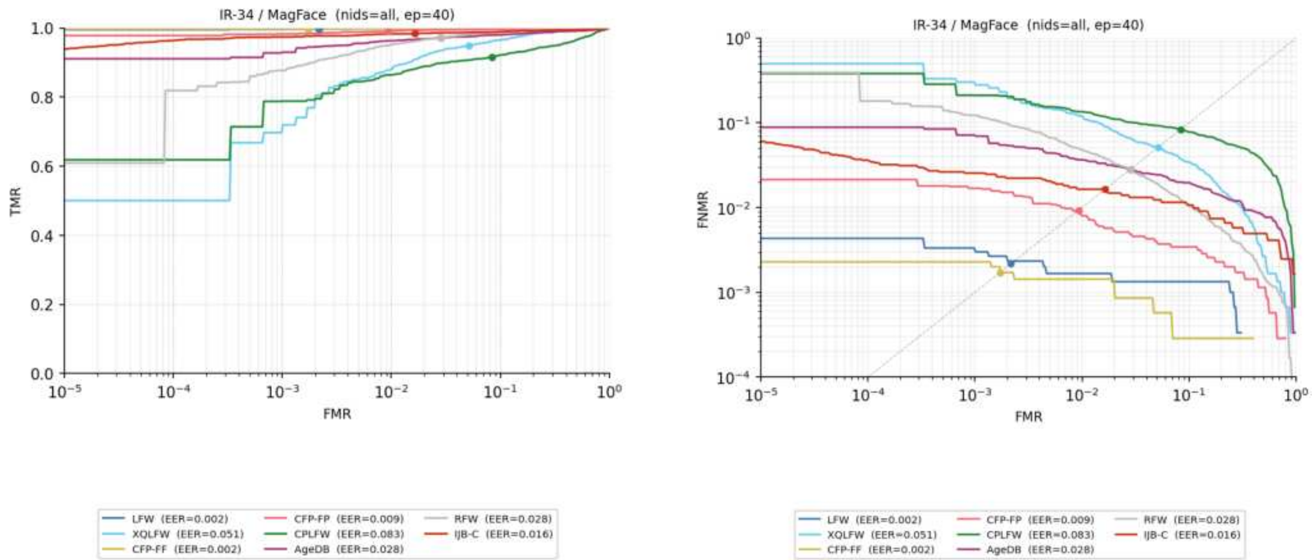


Figure 11. **ROC and DET curves — IResNet-34 / MagFace.** Each curve represents one combination of training-set size (n_{ids}) and training epoch across all evaluation benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-50 / ArcFace

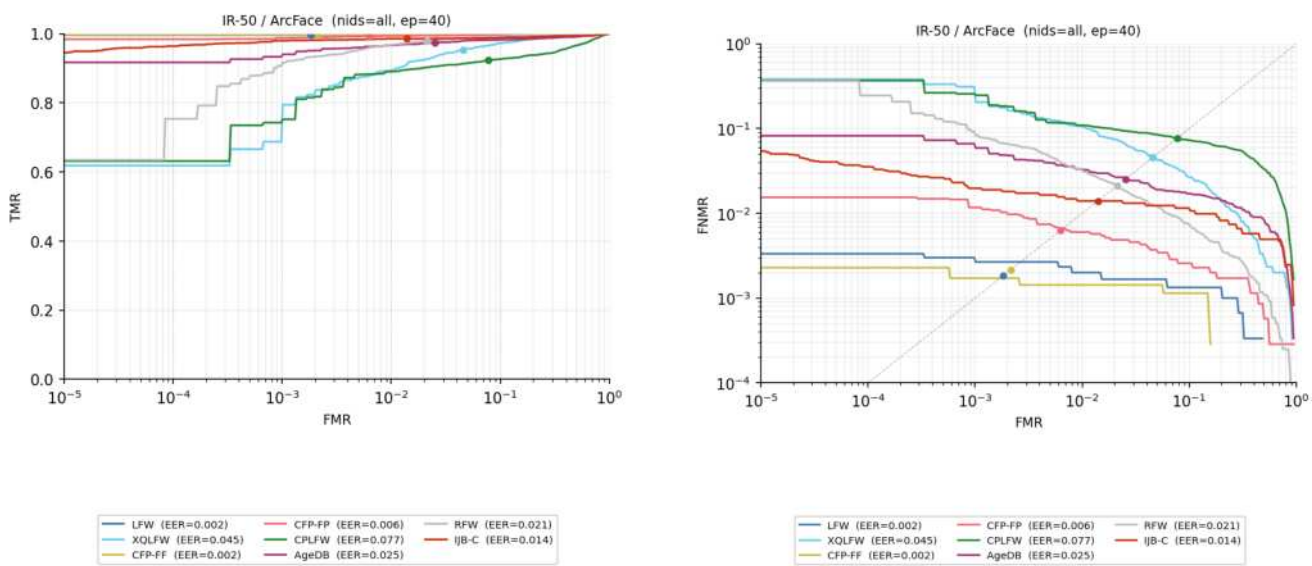


Figure 12. **ROC and DET curves — IResNet-50 / ArcFace.** Each curve represents one combination of training-set size (n_{ids}) and training epoch across all evaluation benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-50 / CosFace

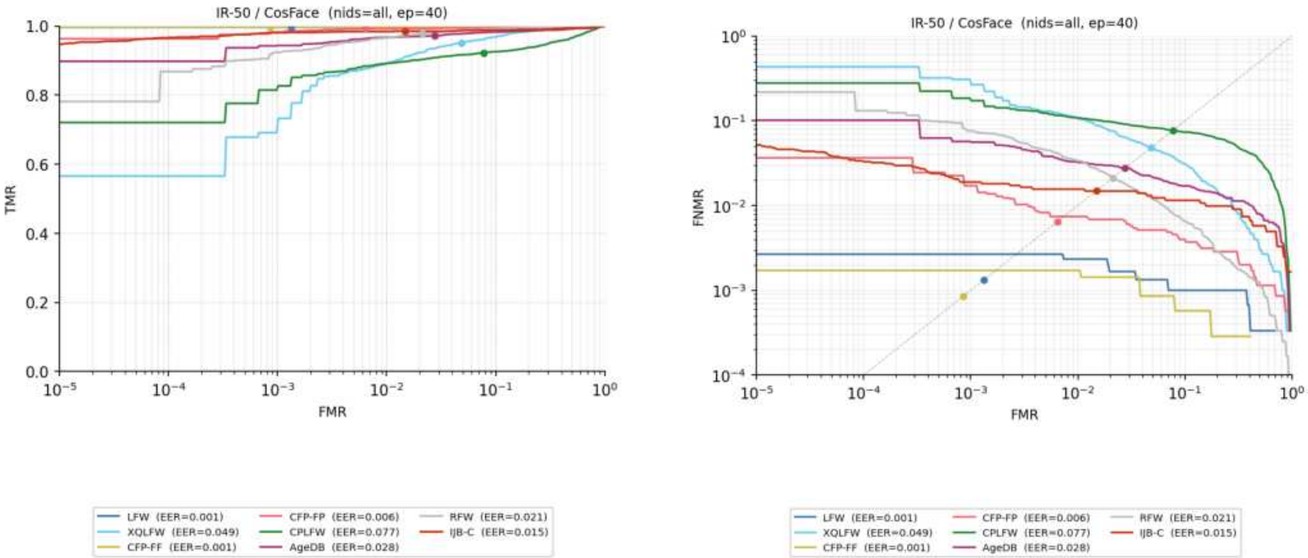


Figure 13. **ROC and DET curves — IResNet-50 / CosFace.** Each curve represents one combination of training-set size (n_{ids}) and training epoch across all evaluation benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-50 / MagFace

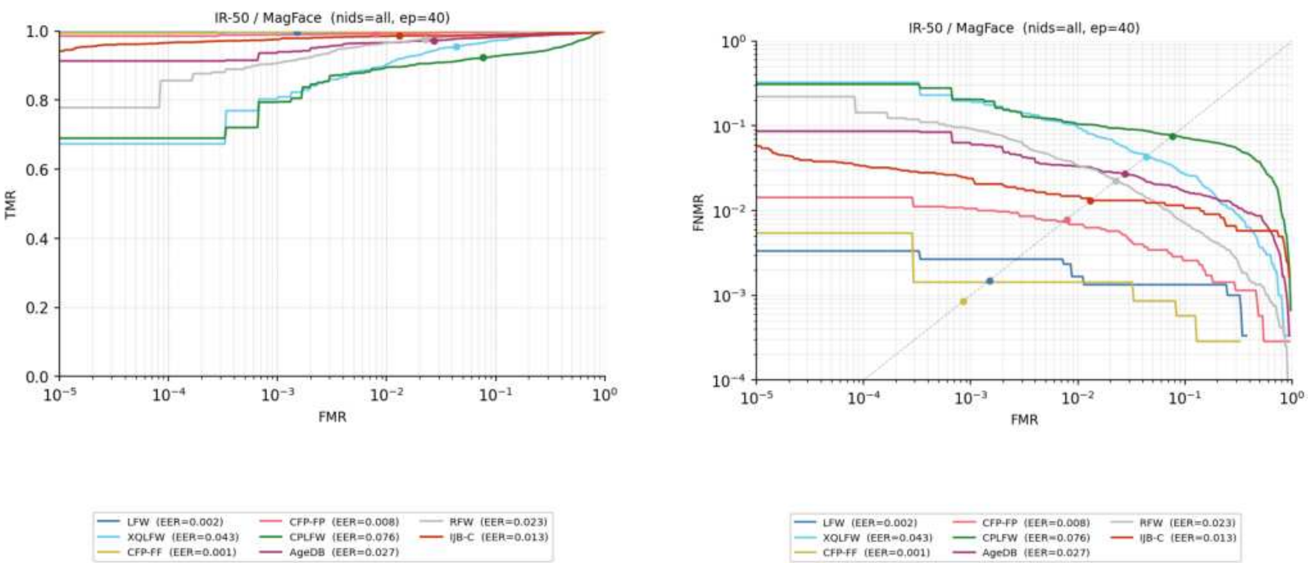


Figure 14. **ROC and DET curves — IResNet-50 / MagFace.** Each curve represents one combination of training-set size (n_{ids}) and training epoch across all evaluation benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-100 / ArcFace

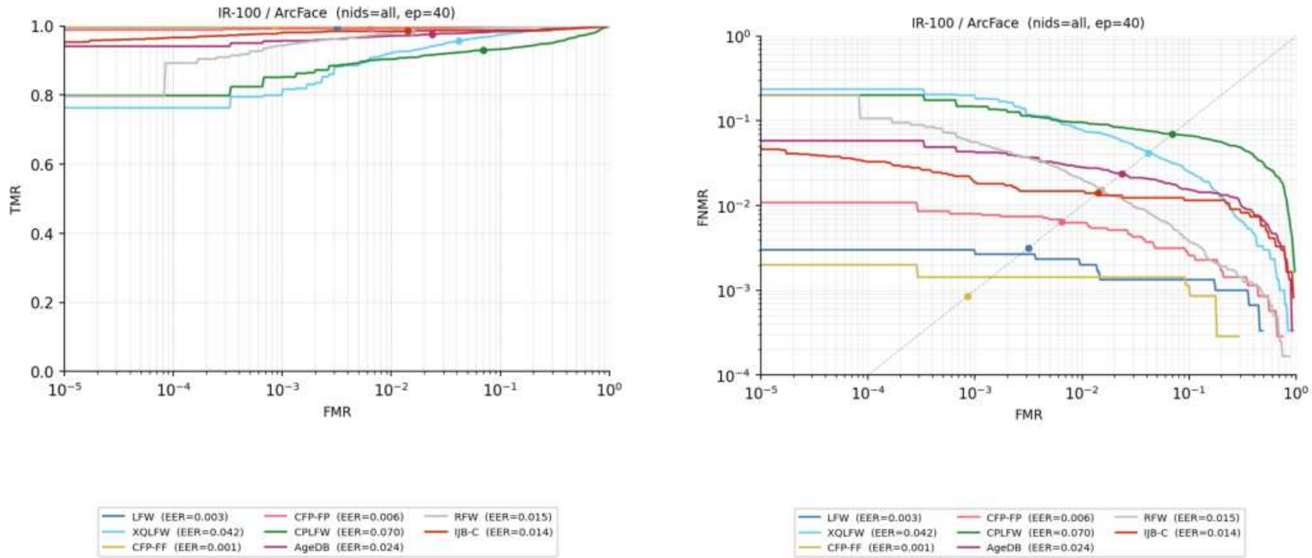


Figure 15. **ROC and DET curves — IResNet-100 / ArcFace.** Each curve represents one combination of training-set size (n_{ids}) and training epoch across all evaluation benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-100 / CosFace

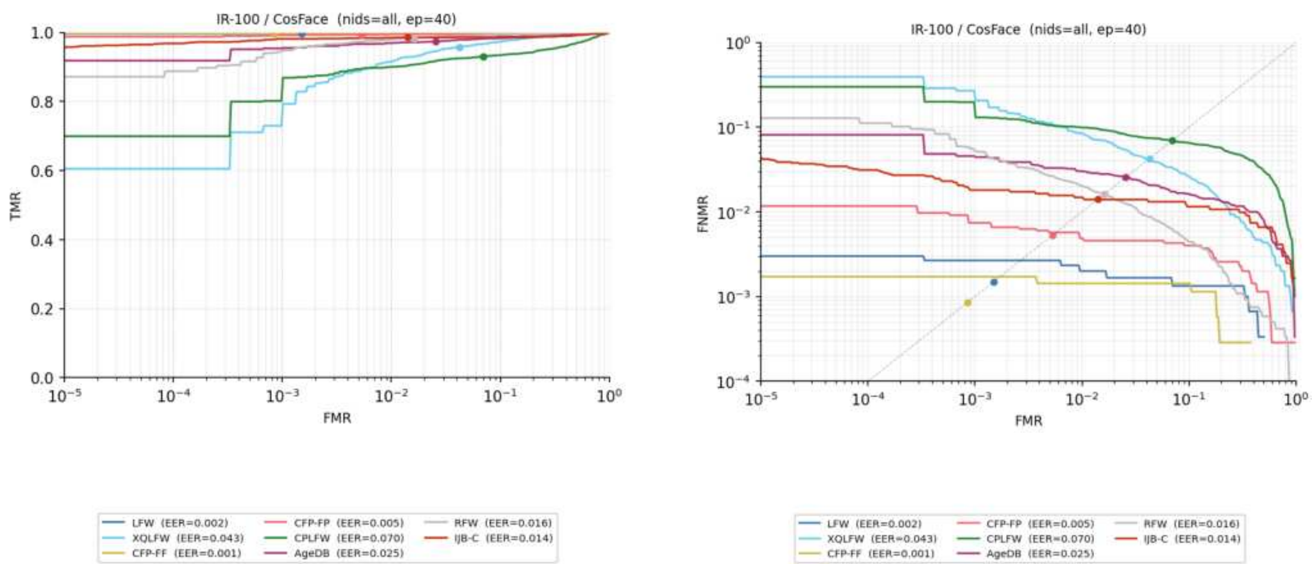


Figure 16. **ROC and DET curves — IResNet-100 / CosFace.** Each curve represents one combination of training-set size (n_{ids}) and training epoch across all evaluation benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-100 / MagFace

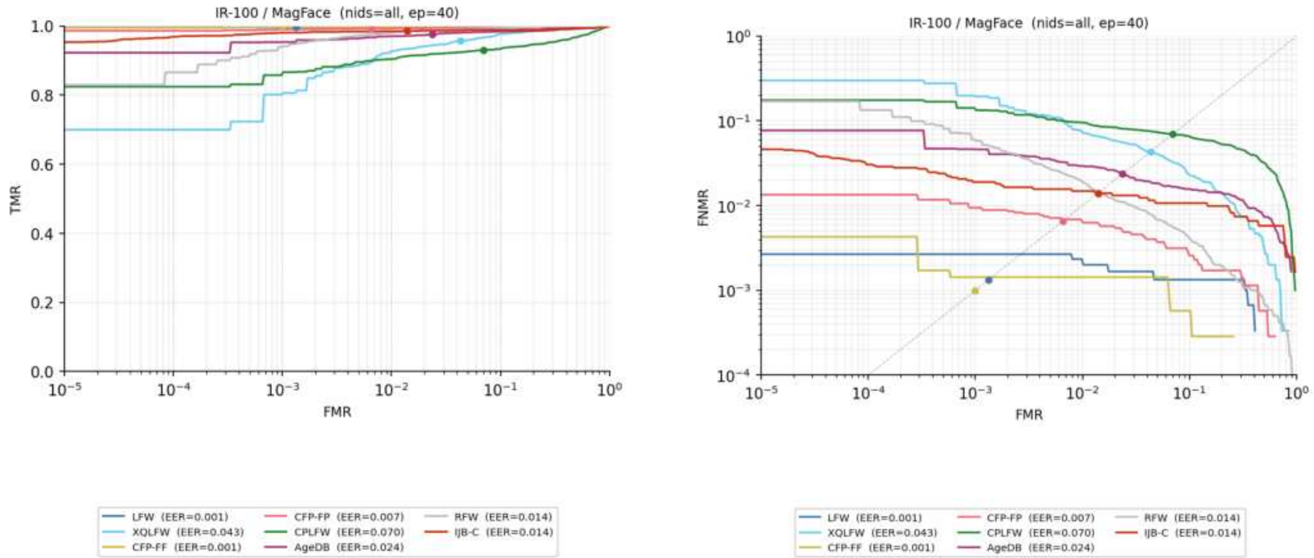


Figure 17. **ROC and DET curves — IResNet-100 / MagFace.** Each curve represents one combination of training-set size (n_{ids}) and training epoch across all evaluation benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

1.3. Grouped by training-set size

AgeDB-30

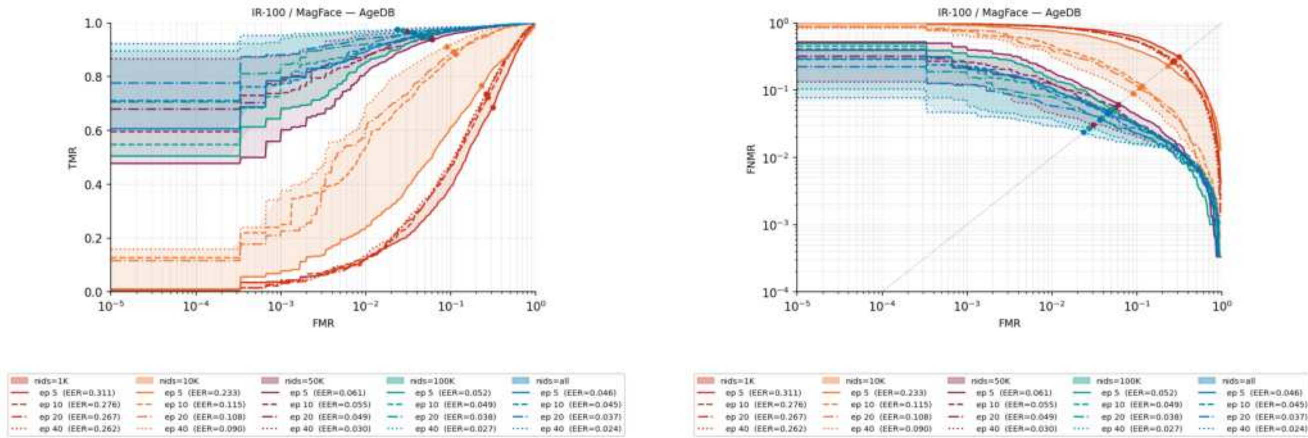


Figure 18. **ROC and DET curves by training-set size — AgeDB-30.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

CFP-FF

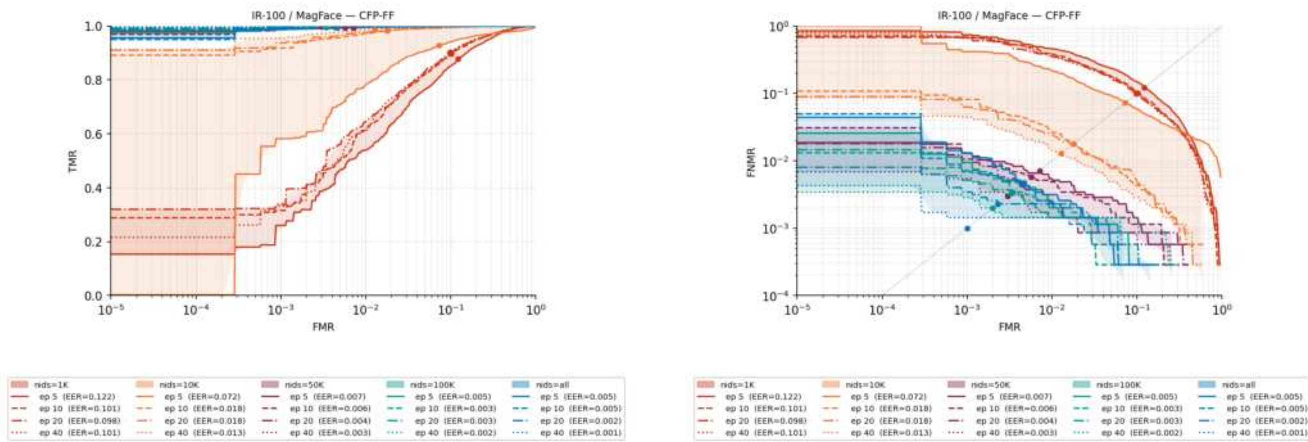


Figure 19. **ROC and DET curves by training-set size — CFP-FF.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

CFP-FP

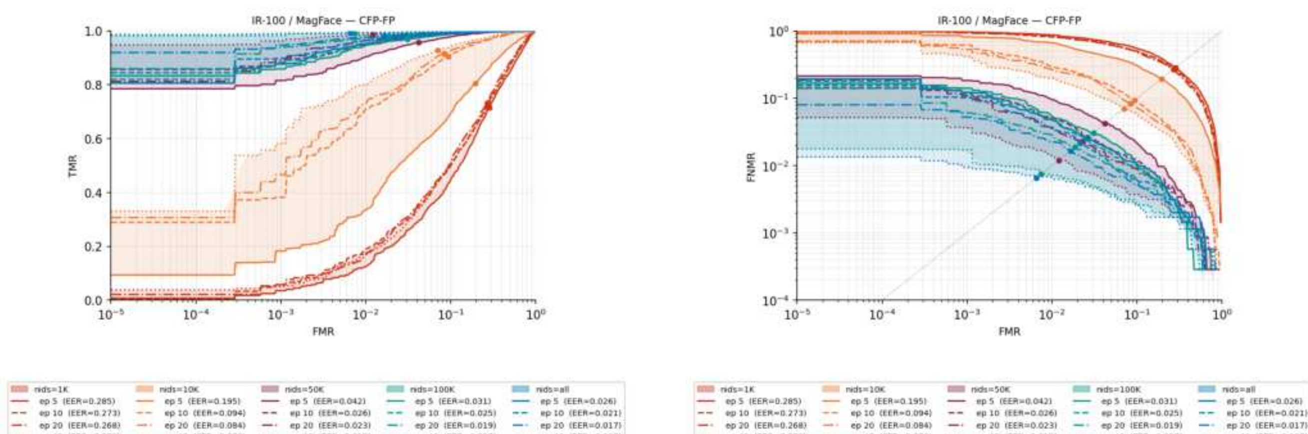


Figure 20. **ROC and DET curves by training-set size — CFP-FP.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

CPLFW

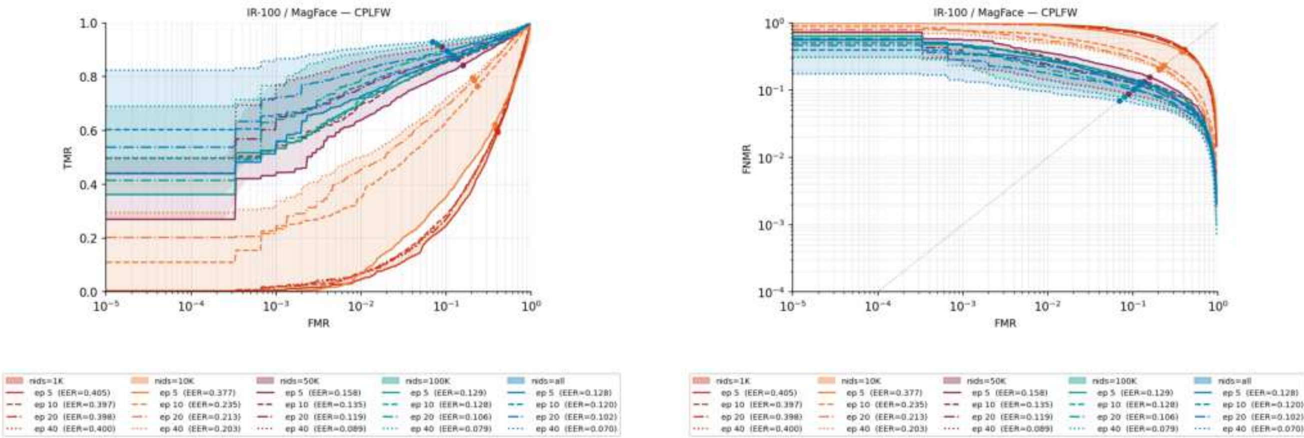


Figure 21. **ROC and DET curves by training-set size — CPLFW.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IJB-C

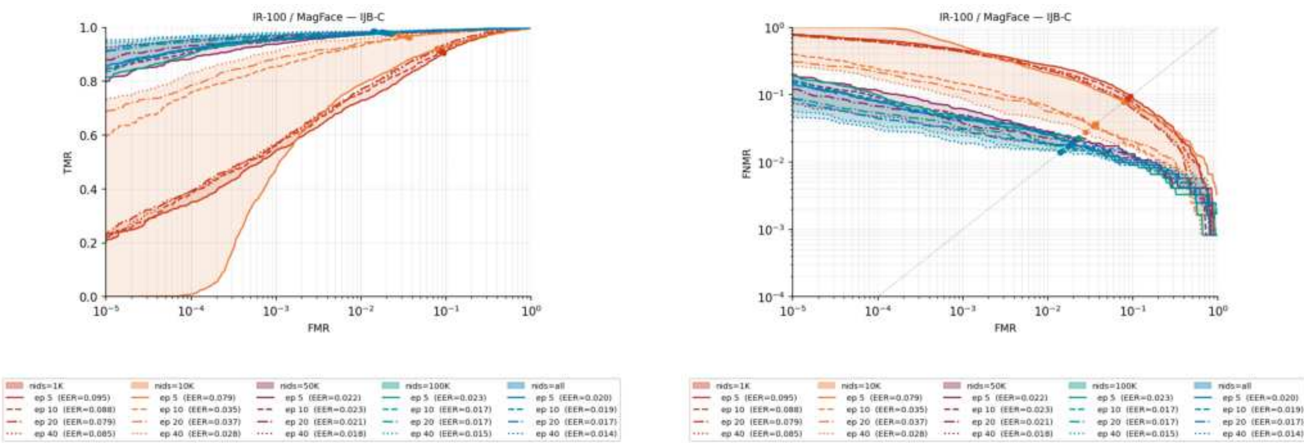


Figure 22. **ROC and DET curves by training-set size — IJB-C.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

LFW

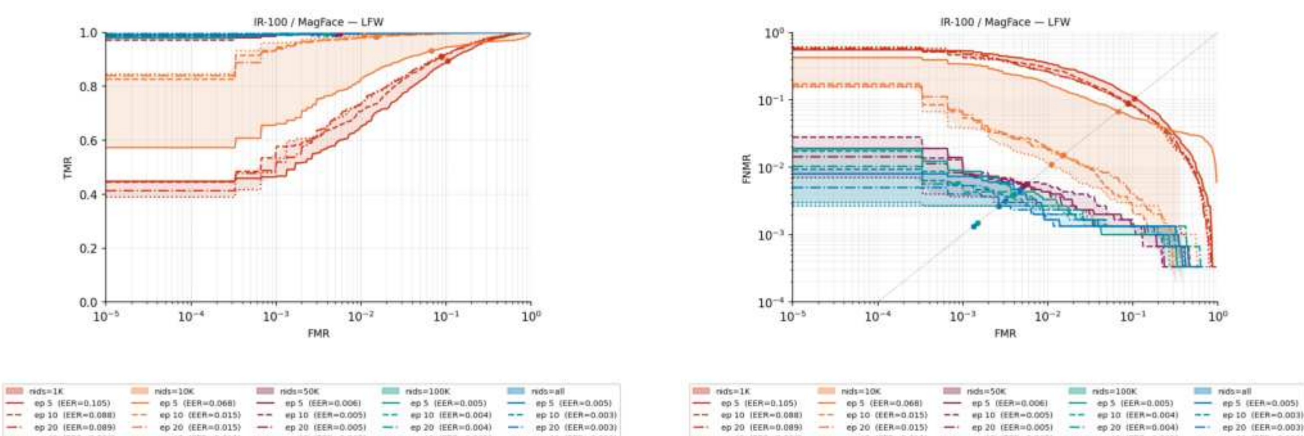


Figure 23. **ROC and DET curves by training-set size — LFW.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

RFW

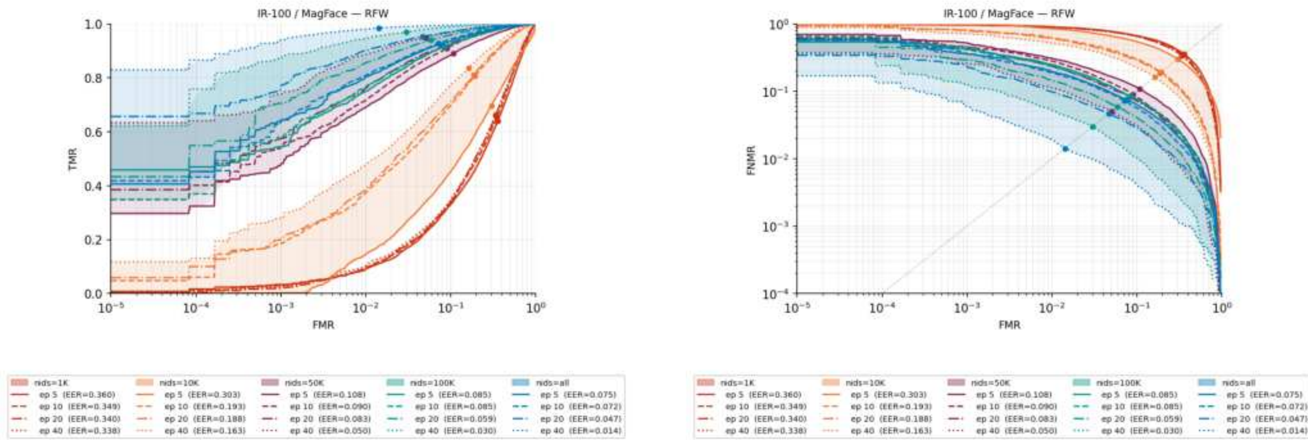


Figure 24. **ROC and DET curves by training-set size — RFW.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K},\ 10\text{K},\ 50\text{K},\ 100\text{K},\ \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

XQLFW

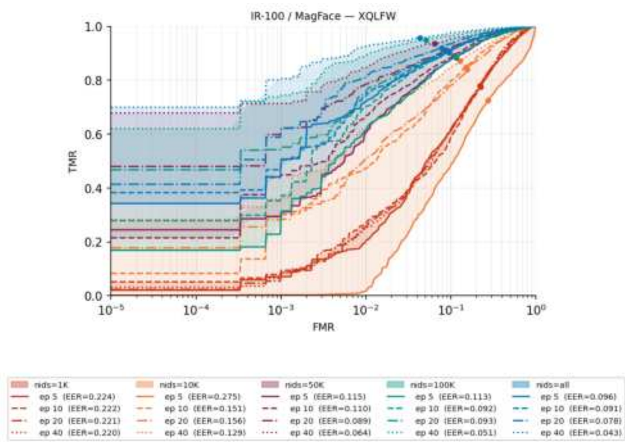


Figure 25. **ROC and DET curves by training-set size — XQLFW.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K},\ 10\text{K},\ 50\text{K},\ 100\text{K},\ \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

2. Membership-classifier ROC and DET curves

Each figure shows the threshold-classifier ROC curve (left) and DET curve (right) for one of the four per-identity geometric statistics. Subsection 2.1 groups by evaluation dataset; Subsection 2.2 groups by backbone and loss head; Subsection 2.3 groups curves by training-set size (n_{ids}) on each dataset, including the same-domain held-out WebFace4M split.

2.1. Grouped by evaluation dataset

2.1.1 PairCos ($\bar{s}$)

AgeDB-30

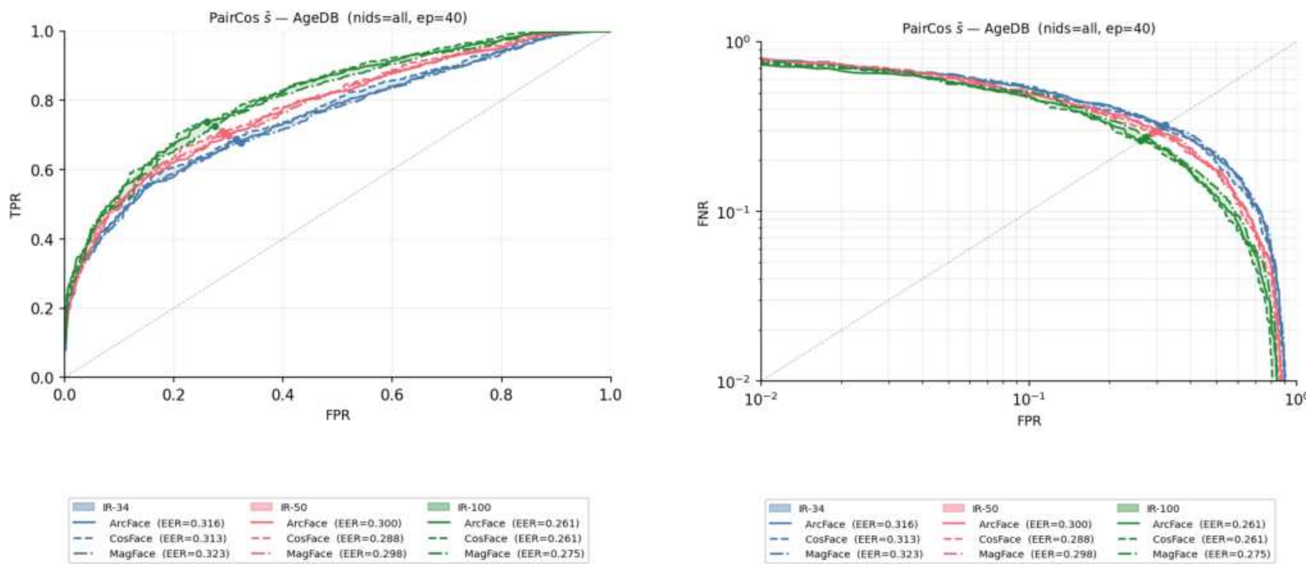


Figure 26. **Threshold-classifier ROC and DET — PairCos ($\bar{s}$), AgeDB-30.** Each curve represents one model configuration (backbone $\times$ loss head $\times$ n_{ids} $\times$ epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

CFP-FF

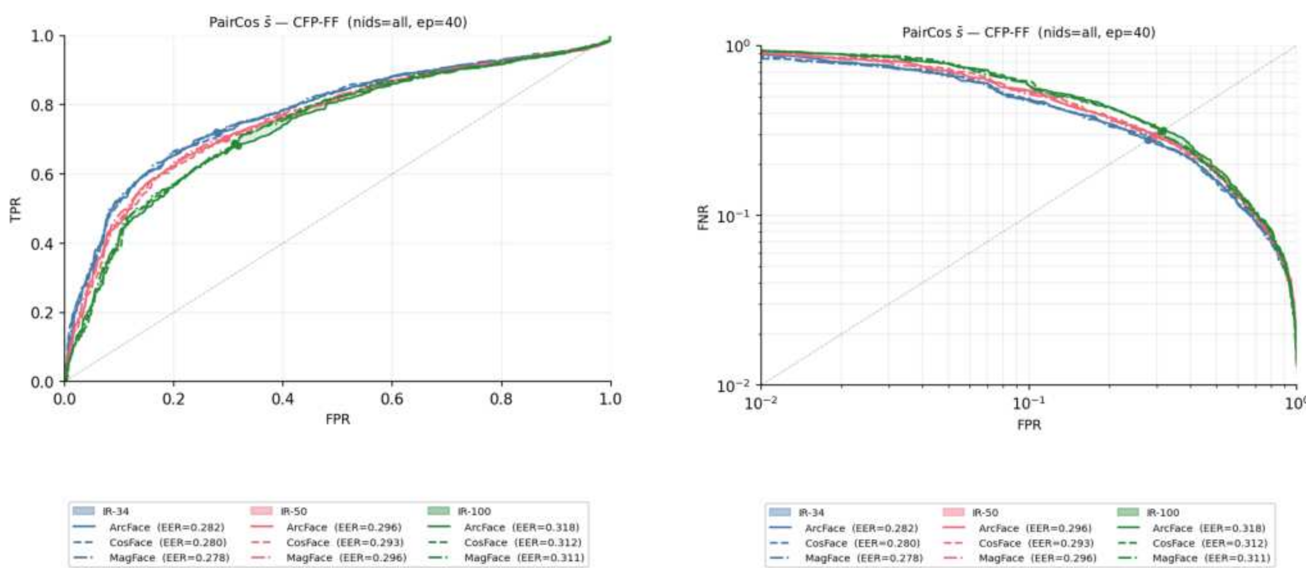


Figure 27. **Threshold-classifier ROC and DET — PairCos ($\bar{s}$), CFP-FF.** Each curve represents one model configuration (backbone $\times$ loss head $\times$ n_{ids} $\times$ epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

CFP-FP

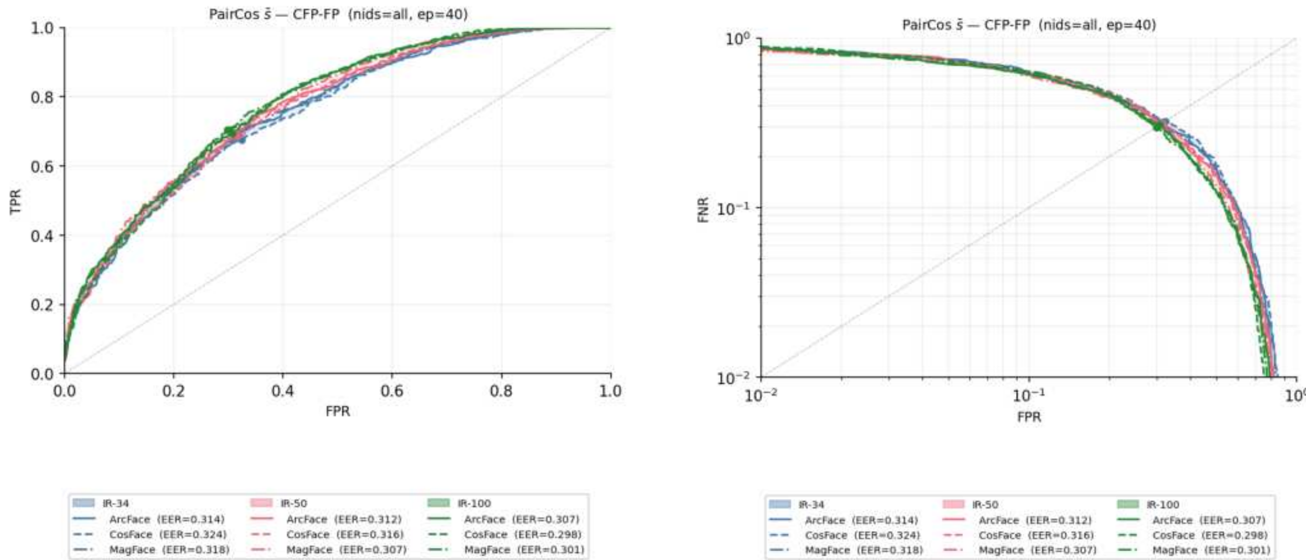


Figure 28. **Threshold-classifier ROC and DET — PairCos ($\bar{s}$), CFP-FP.** Each curve represents one model configuration (backbone $\times$ loss head $\times$ n_{ids} $\times$ epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

CPLFW

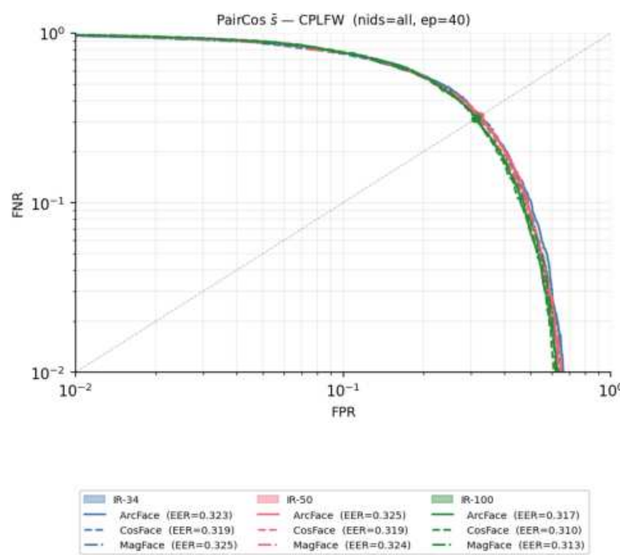


Figure 29. **Threshold-classifier ROC and DET — PairCos ($\bar{s}$), CPLFW.** Each curve represents one model configuration (backbone $\times$ loss head $\times$ n_{ids} $\times$ epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

LFW

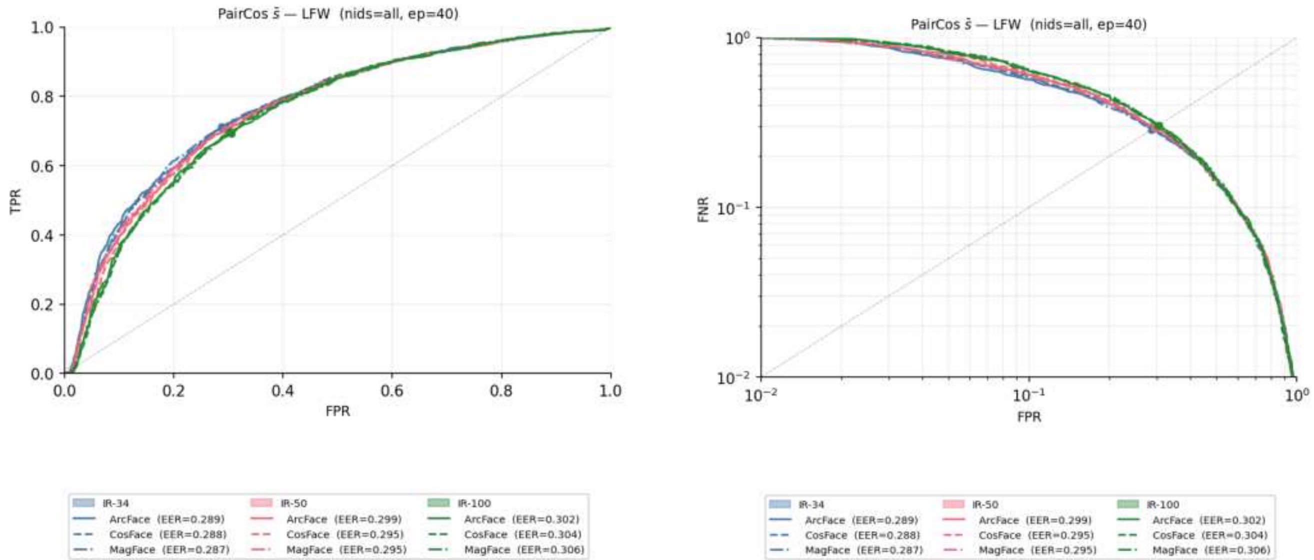


Figure 30. **Threshold-classifier ROC and DET — PairCos ($\bar{s}$), LFW.** Each curve represents one model configuration (backbone × loss head × n_{ids} × epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

RFW

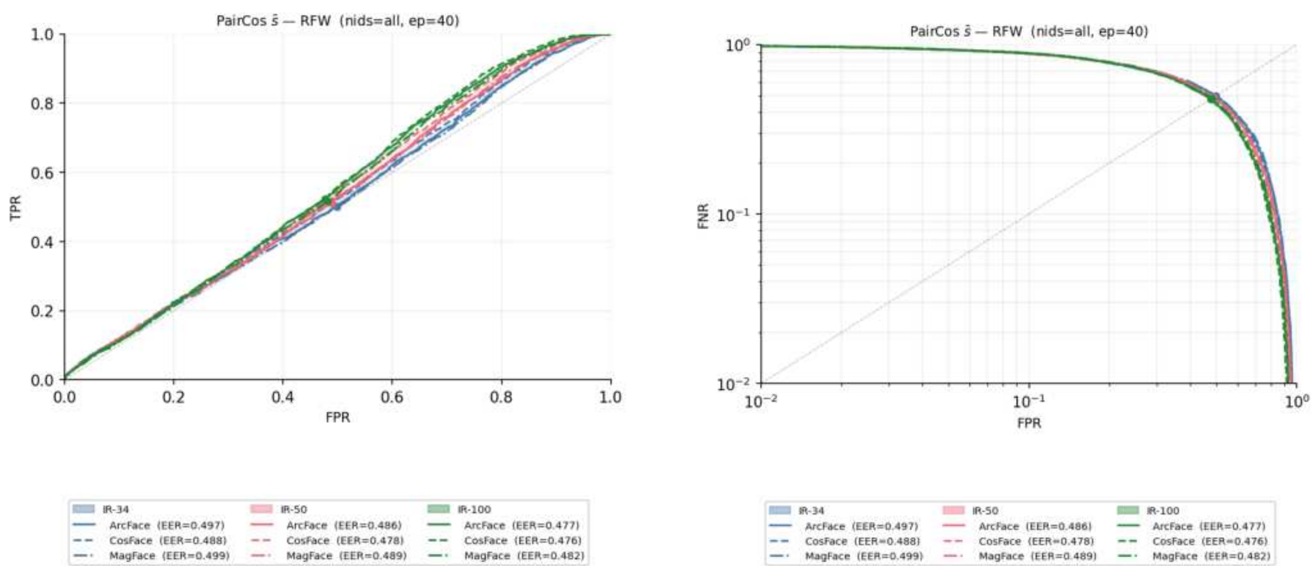


Figure 31. **Threshold-classifier ROC and DET — PairCos ($\bar{s}$), RFW.** Each curve represents one model configuration (backbone × loss head × n_{ids} × epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

XQLFW

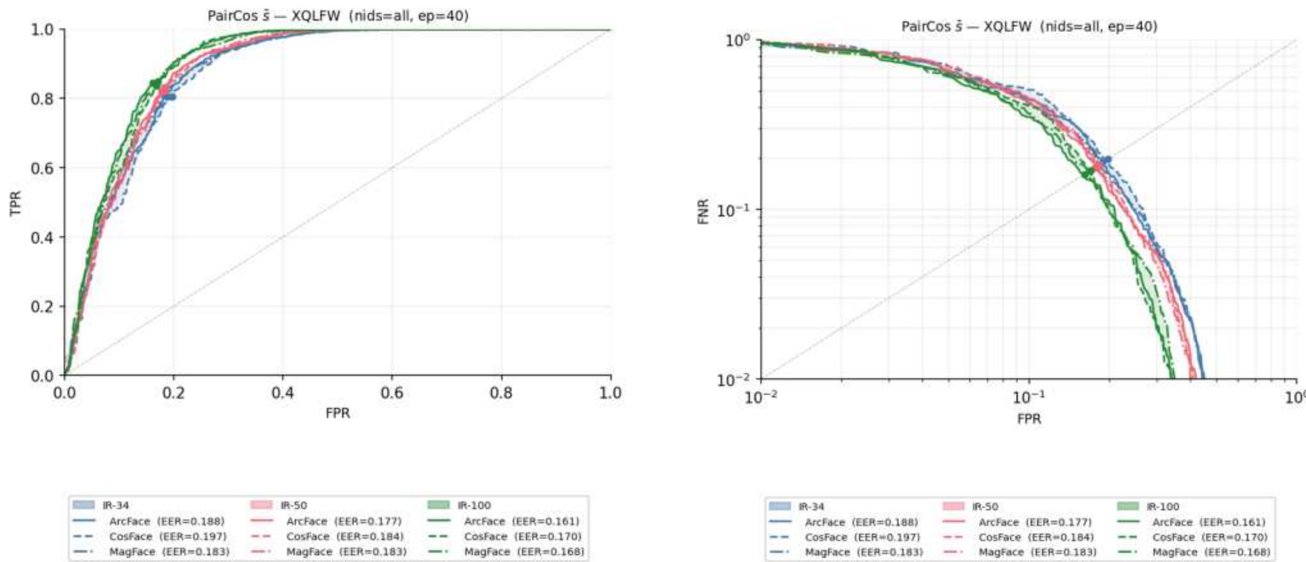


Figure 32. **Threshold-classifier ROC and DET — PairCos ($\bar{s}$), XQLFW.** Each curve represents one model configuration (backbone $\times$ loss head $\times$ n_{ids} $\times$ epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

2.1.2 vMF-κ

AgeDB-30

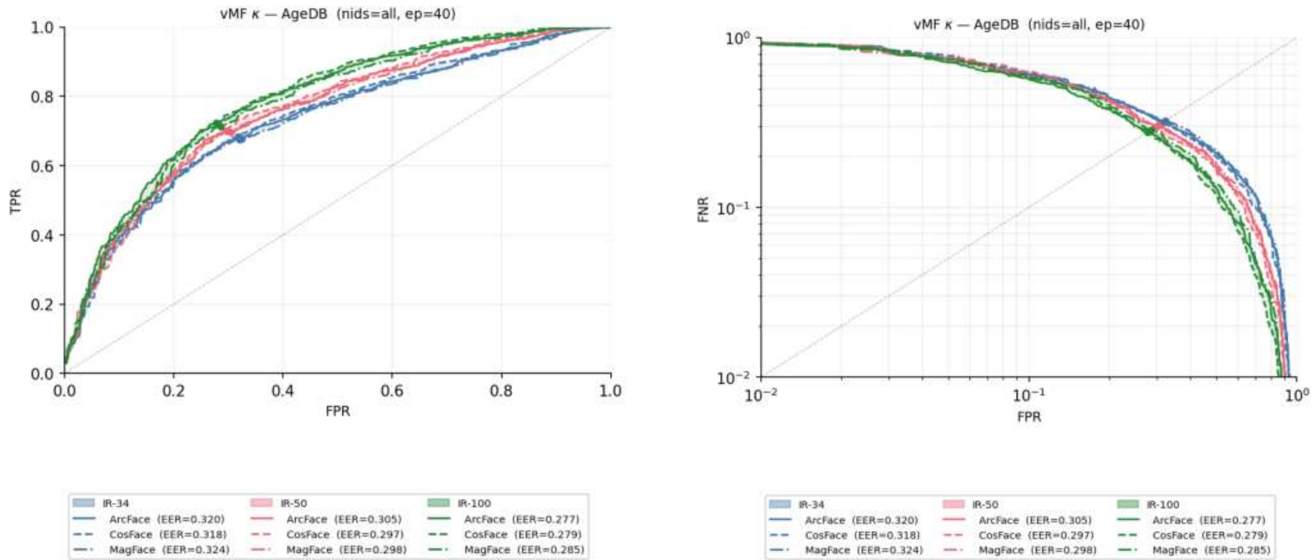


Figure 33. **Threshold-classifier ROC and DET — vMF-κ, AgeDB-30.** Each curve represents one model configuration (backbone $\times$ loss head $\times$ n_{ids} $\times$ epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

CFP-FF

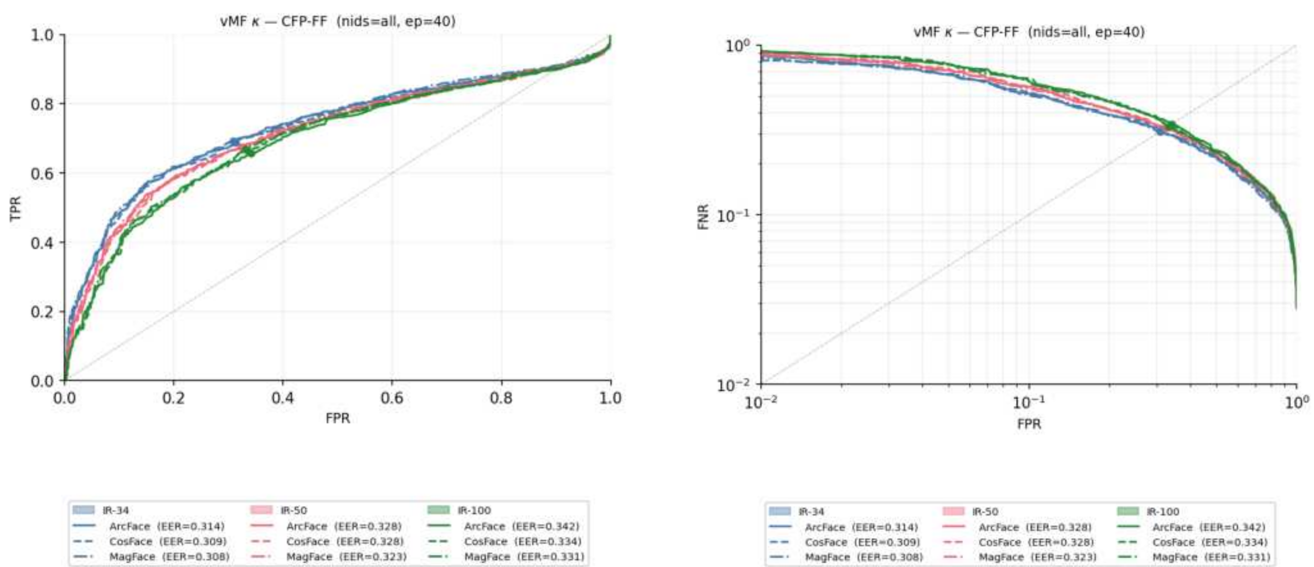


Figure 34. **Threshold-classifier ROC and DET — vMF-κ, CFP-FF.** Each curve represents one model configuration (backbone $\times$ loss head $\times$ n_{ids} $\times$ epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

CFP-FP

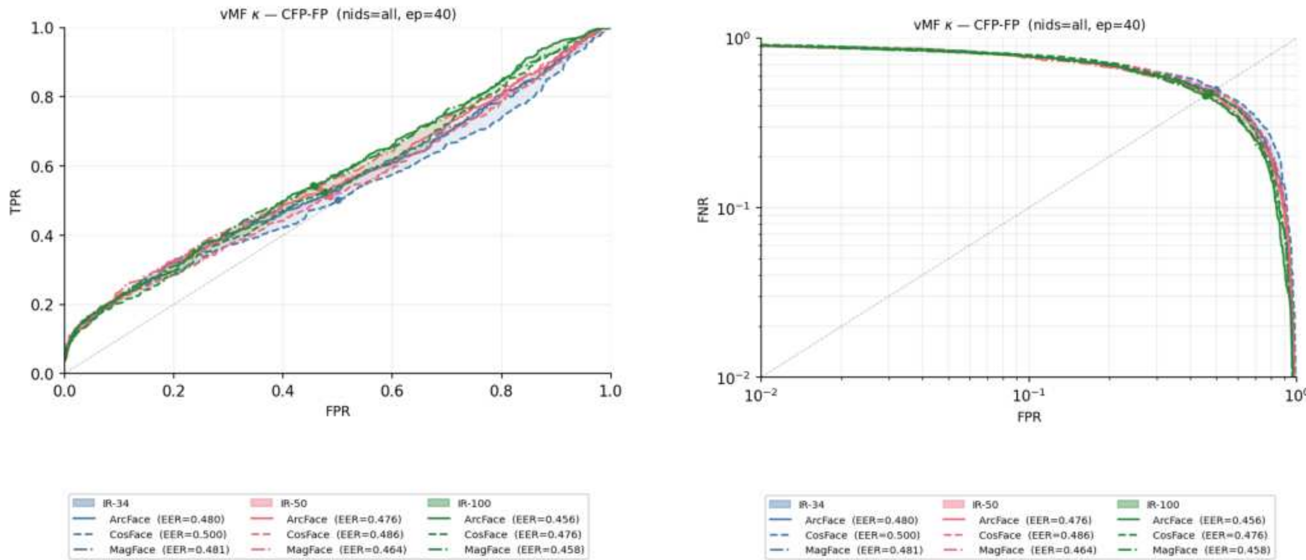


Figure 35. **Threshold-classifier ROC and DET — vMF-κ, CFP-FP.** Each curve represents one model configuration (backbone $\times$ loss head $\times$ n_{ids} $\times$ epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

CPLFW

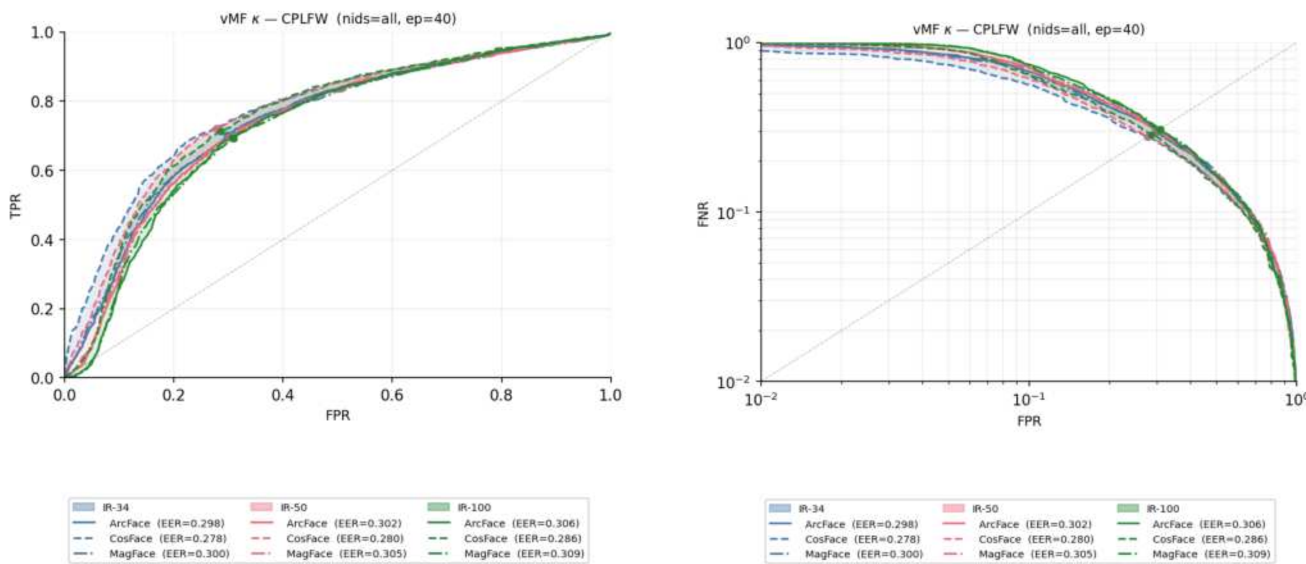


Figure 36. **Threshold-classifier ROC and DET — vMF-κ, CPLFW.** Each curve represents one model configuration (backbone $\times$ loss head $\times$ n_{ids} $\times$ epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

LFW

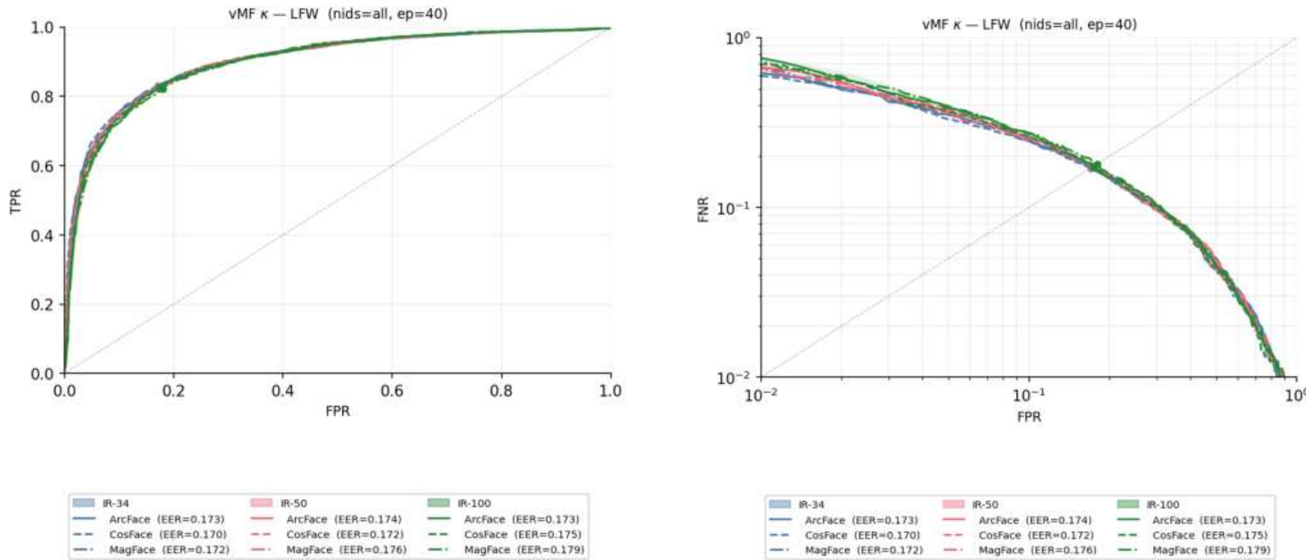


Figure 37. **Threshold-classifier ROC and DET — vMF-κ, LFW.** Each curve represents one model configuration (backbone $\times$ loss head $\times$ n_{ids} $\times$ epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

RFW

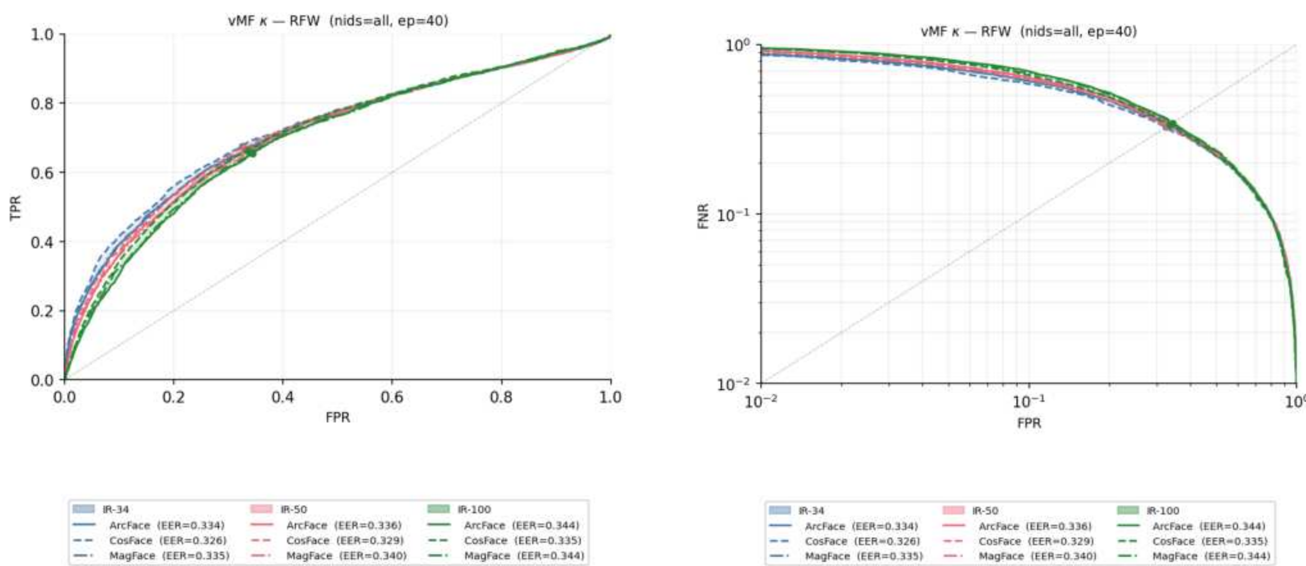


Figure 38. **Threshold-classifier ROC and DET — vMF-κ, RFW.** Each curve represents one model configuration (backbone $\times$ loss head $\times$ n_{ids} $\times$ epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

XQLFW

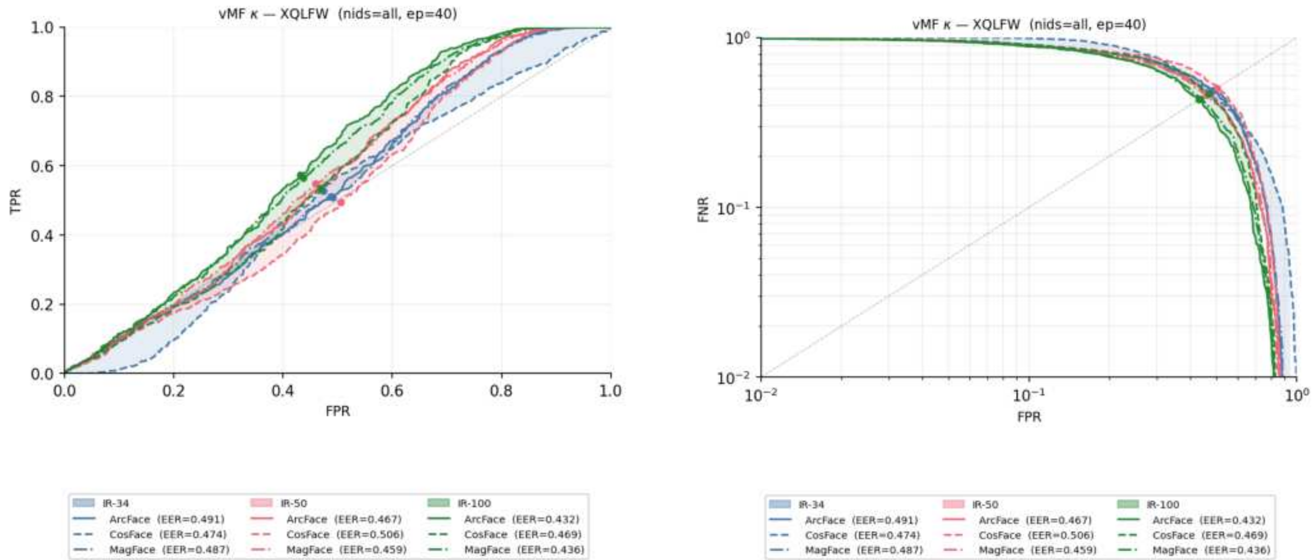


Figure 39. **Threshold-classifier ROC and DET — vMF-κ, XQLFW.** Each curve represents one model configuration (backbone $\times$ loss head $\times$ n_{ids} $\times$ epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

2.1.3 PenLogit

AgeDB-30

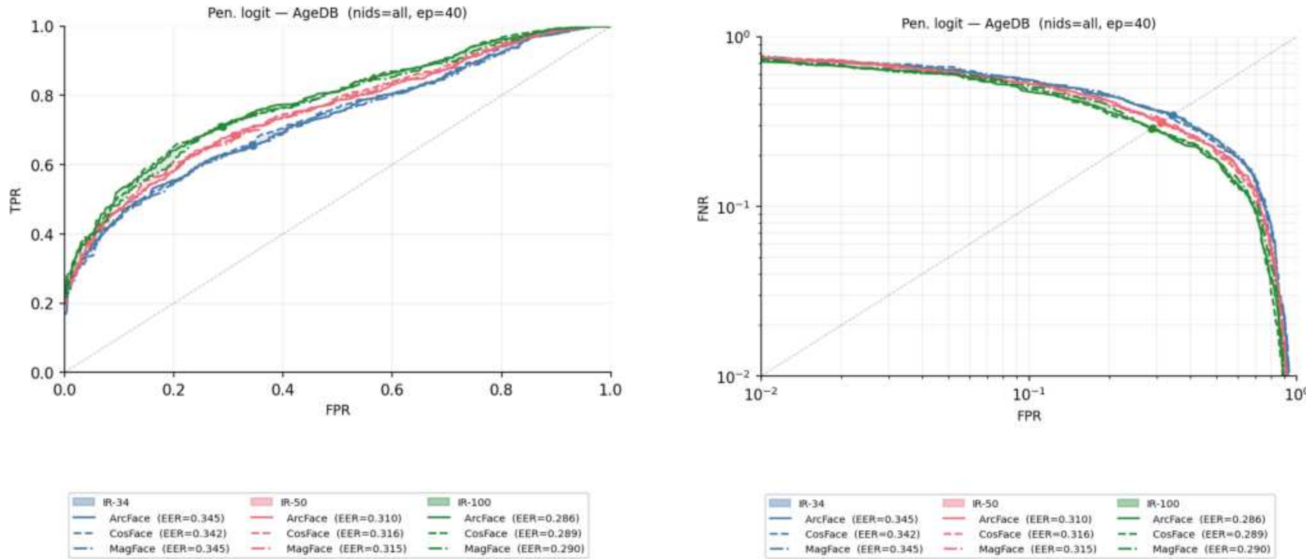


Figure 40. **Threshold-classifier ROC and DET — PenLogit, AgeDB-30.** Each curve represents one model configuration (backbone $\times$ loss head $\times$ n_{ids} $\times$ epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

CFP-FF

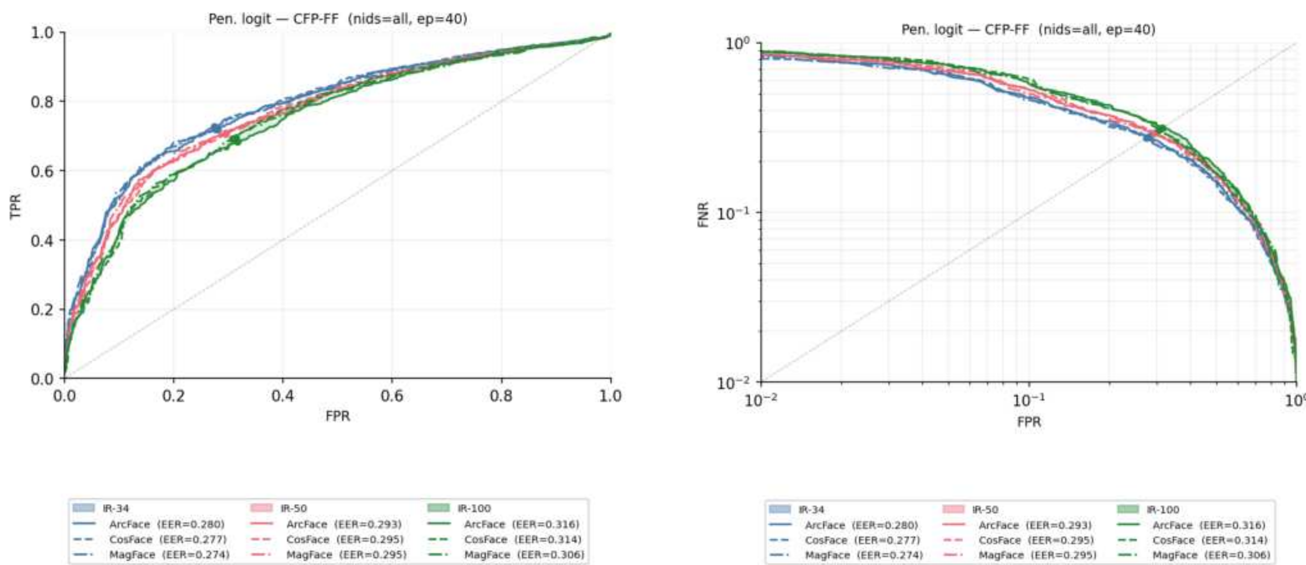


Figure 41. **Threshold-classifier ROC and DET — PenLogit, CFP-FF.** Each curve represents one model configuration (backbone $\times$ loss head $\times$ n_{ids} $\times$ epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

CFP-FP

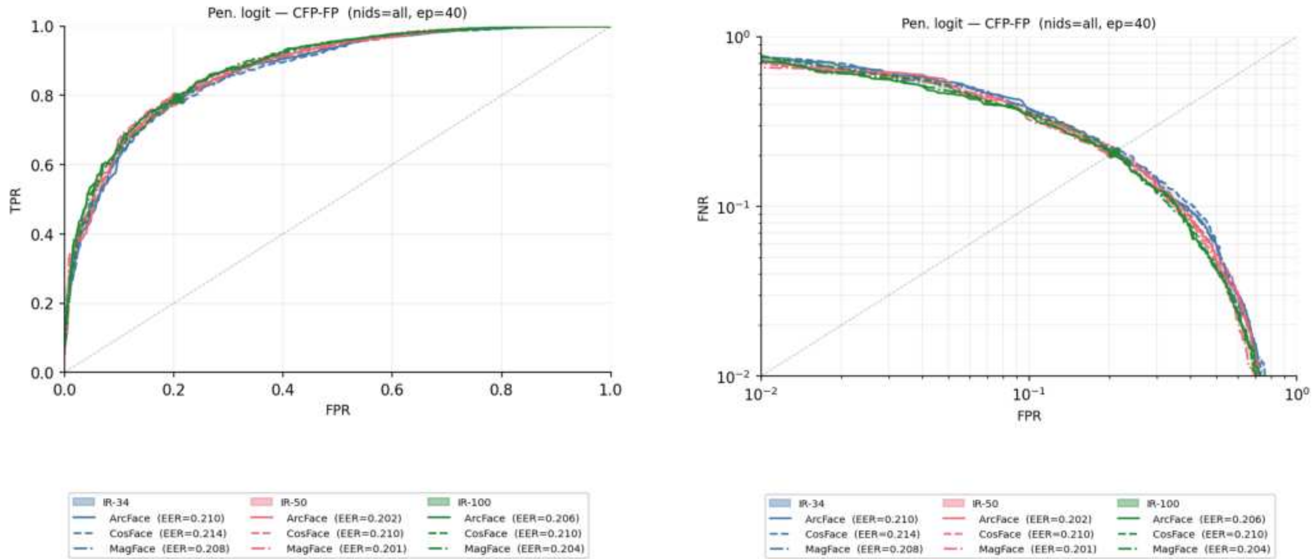


Figure 42. **Threshold-classifier ROC and DET — PenLogit, CFP-FP.** Each curve represents one model configuration (backbone $\times$ loss head $\times$ n_{ids} $\times$ epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

CPLFW

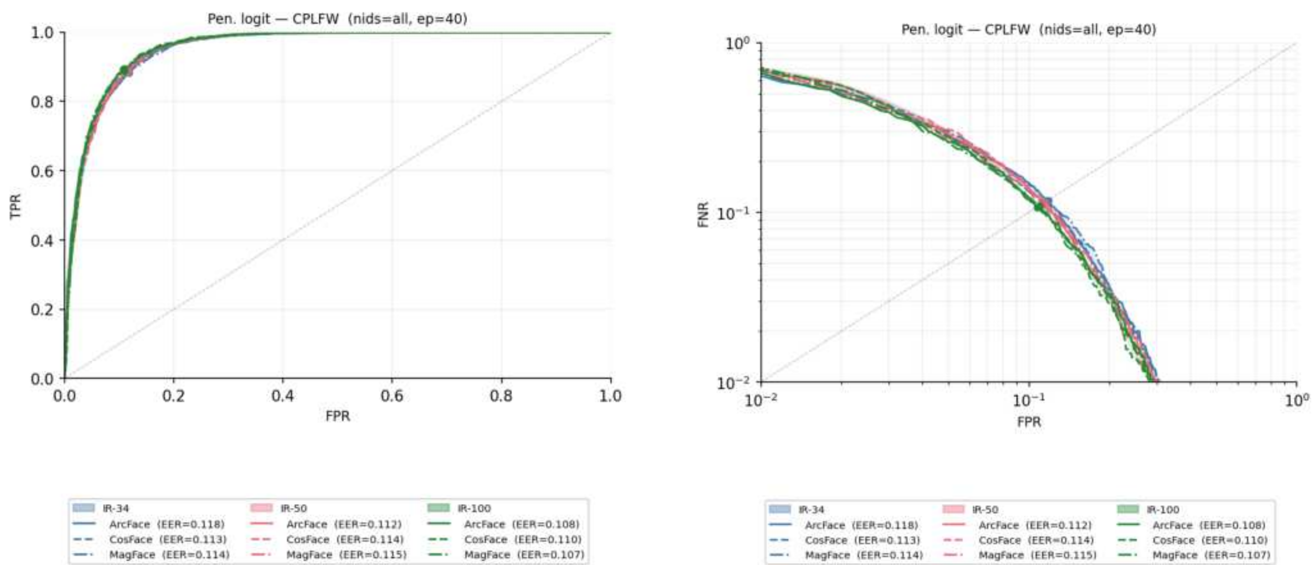


Figure 43. **Threshold-classifier ROC and DET — PenLogit, CPLFW.** Each curve represents one model configuration (backbone $\times$ loss head $\times$ n_{ids} $\times$ epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

LFW

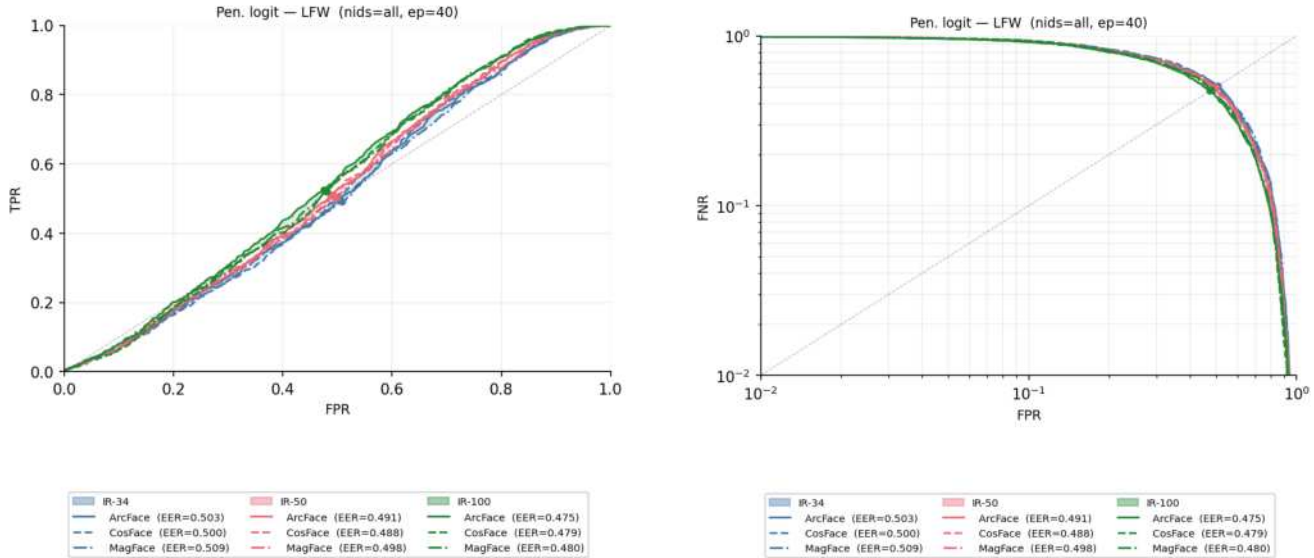


Figure 44. **Threshold-classifier ROC and DET — PenLogit, LFW.** Each curve represents one model configuration (backbone $\times$ loss head $\times$ n_{ids} $\times$ epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

RFW

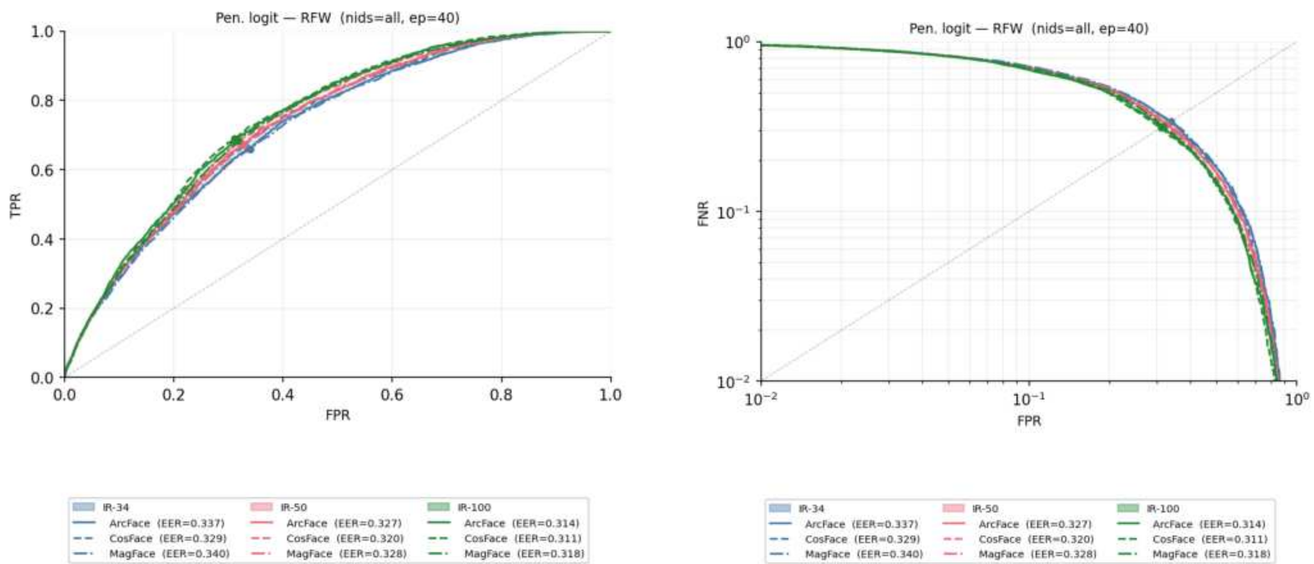


Figure 45. **Threshold-classifier ROC and DET — PenLogit, RFW.** Each curve represents one model configuration (backbone $\times$ loss head $\times$ n_{ids} $\times$ epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

XQLFW

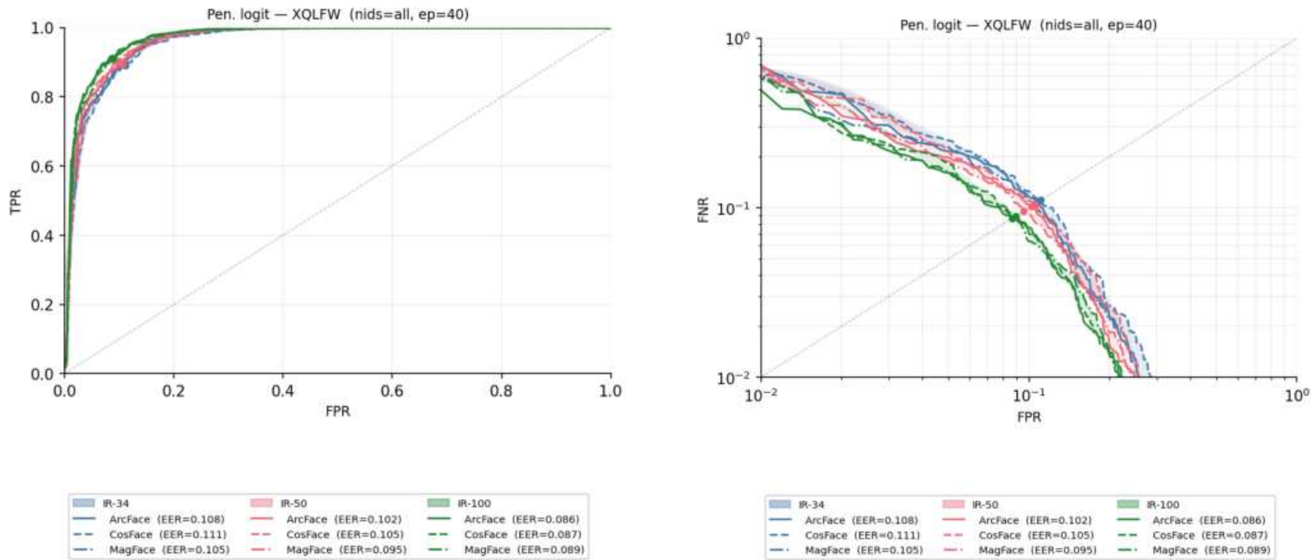


Figure 46. **Threshold-classifier ROC and DET — PenLogit, XQLFW.** Each curve represents one model configuration (backbone × loss head × n_{ids} × epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

2.1.4 ProtoCE

AgeDB-30

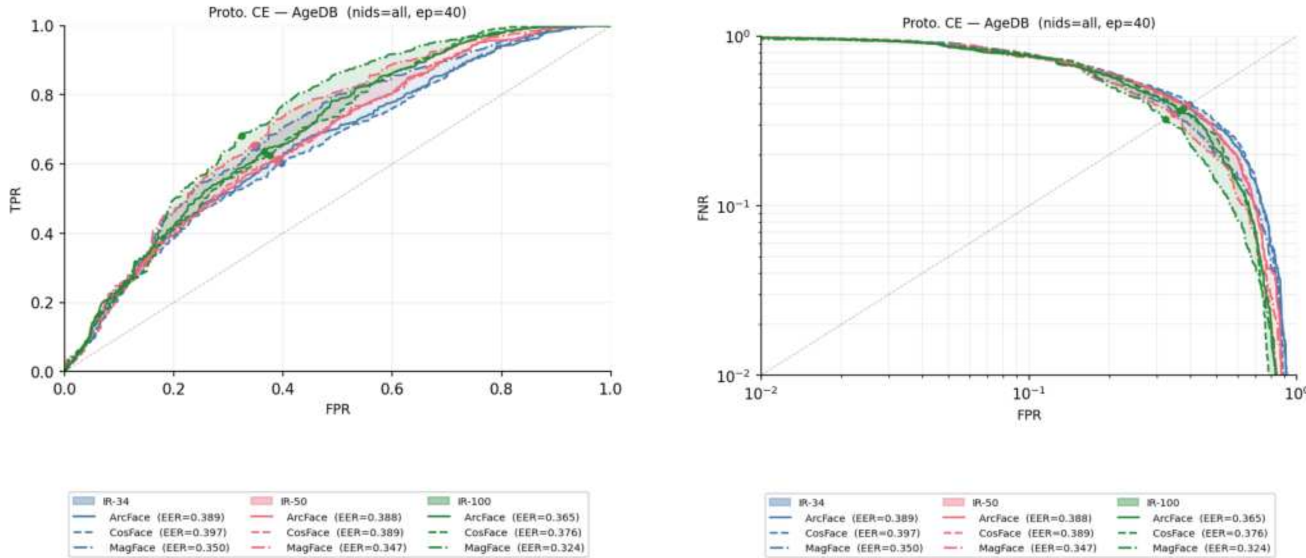


Figure 47. **Threshold-classifier ROC and DET — ProtoCE, AgeDB-30.** Each curve represents one model configuration (backbone $\times$ loss head $\times$ n_{ids} $\times$ epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

CFP-FF

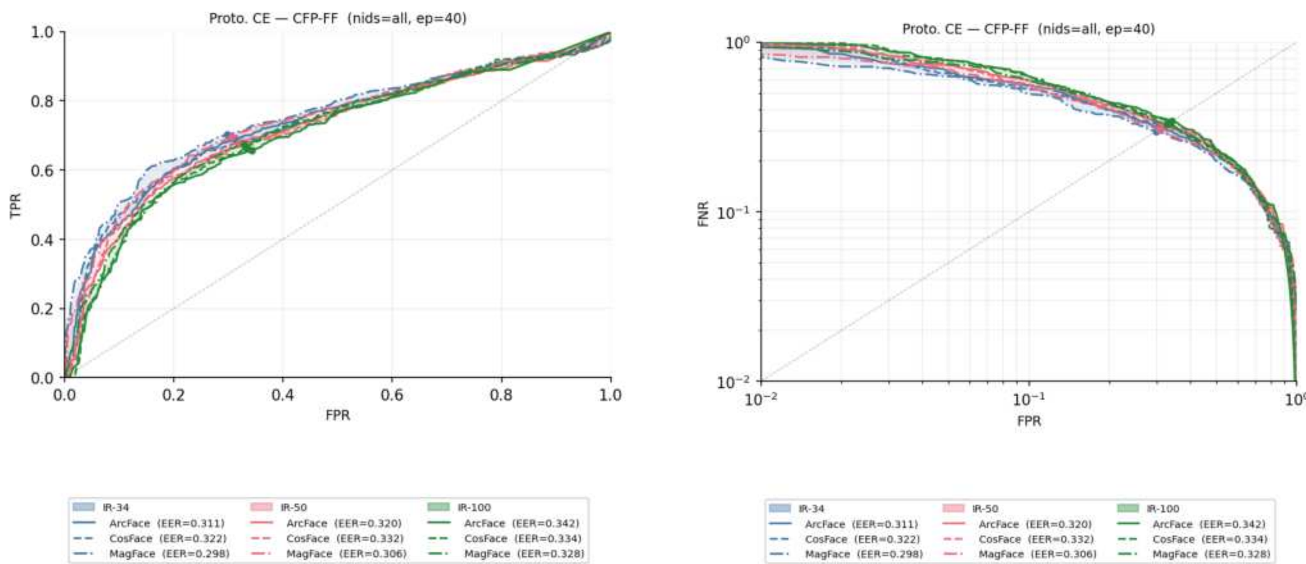


Figure 48. **Threshold-classifier ROC and DET — ProtoCE, CFP-FF.** Each curve represents one model configuration (backbone $\times$ loss head $\times$ n_{ids} $\times$ epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

CFP-FP

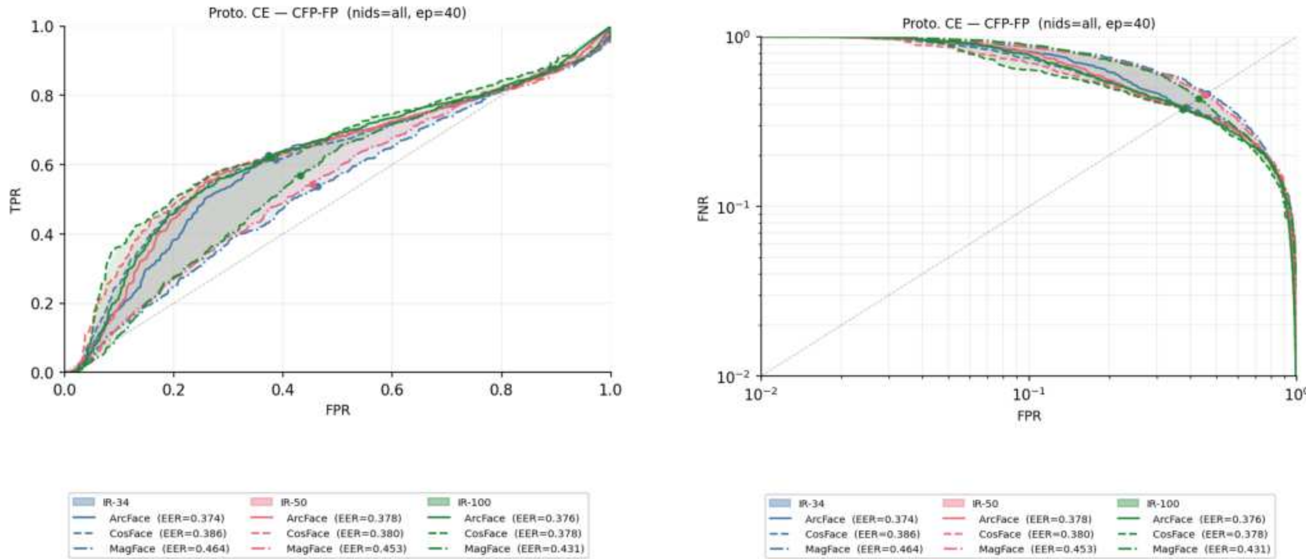


Figure 49. **Threshold-classifier ROC and DET — ProtoCE, CFP-FP.** Each curve represents one model configuration (backbone $\times$ loss head $\times$ n_{ids} $\times$ epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

CPLFW

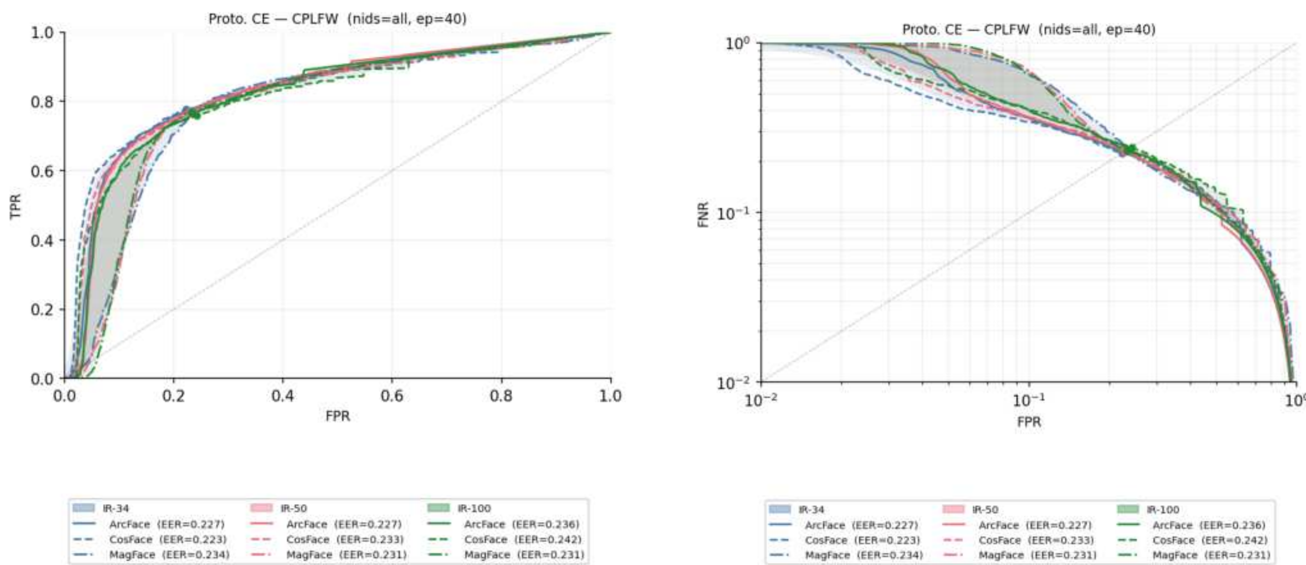


Figure 50. **Threshold-classifier ROC and DET — ProtoCE, CPLFW.** Each curve represents one model configuration (backbone $\times$ loss head $\times$ n_{ids} $\times$ epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

LFW

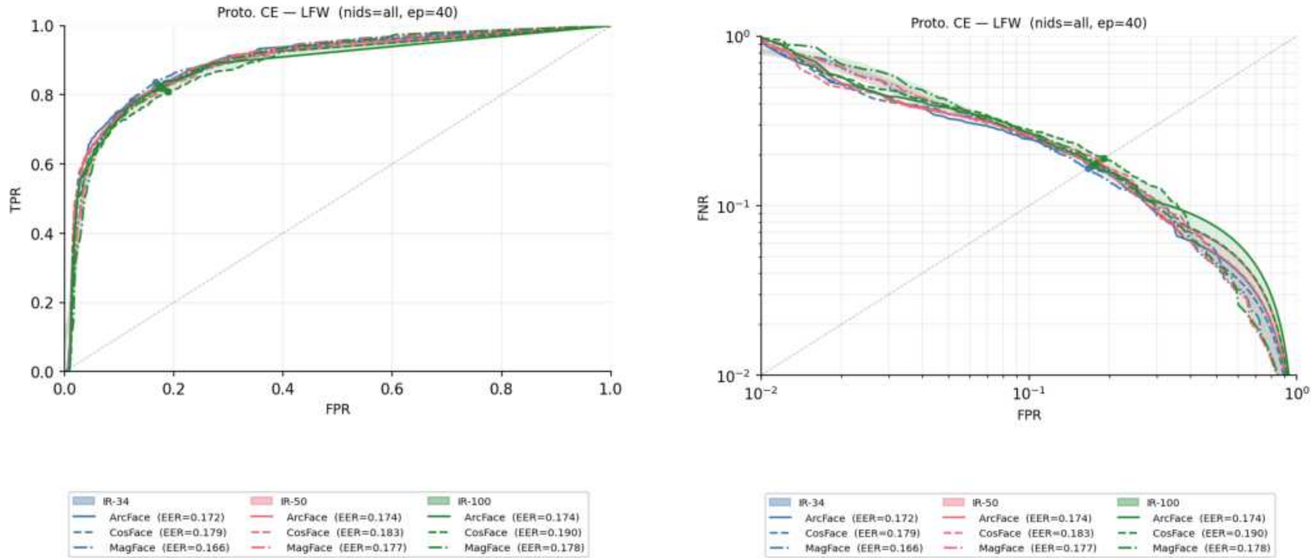


Figure 51. **Threshold-classifier ROC and DET — ProtoCE, LFW.** Each curve represents one model configuration (backbone × loss head × n_{ids} × epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

RFW

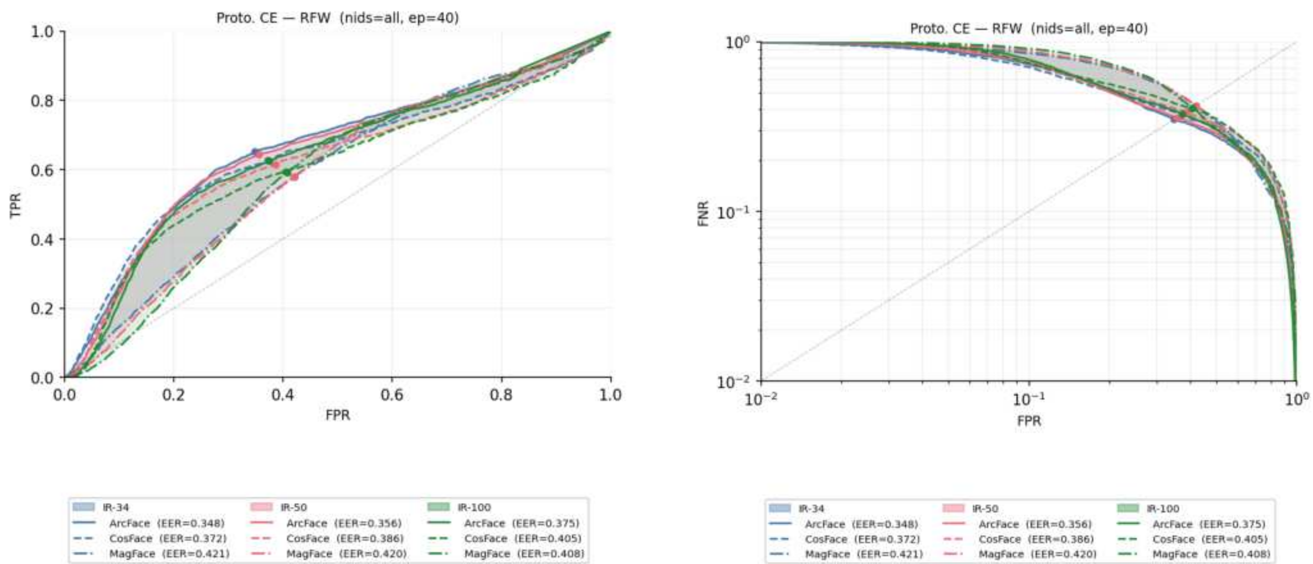


Figure 52. **Threshold-classifier ROC and DET — ProtoCE, RFW.** Each curve represents one model configuration (backbone × loss head × n_{ids} × epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

XQLFW

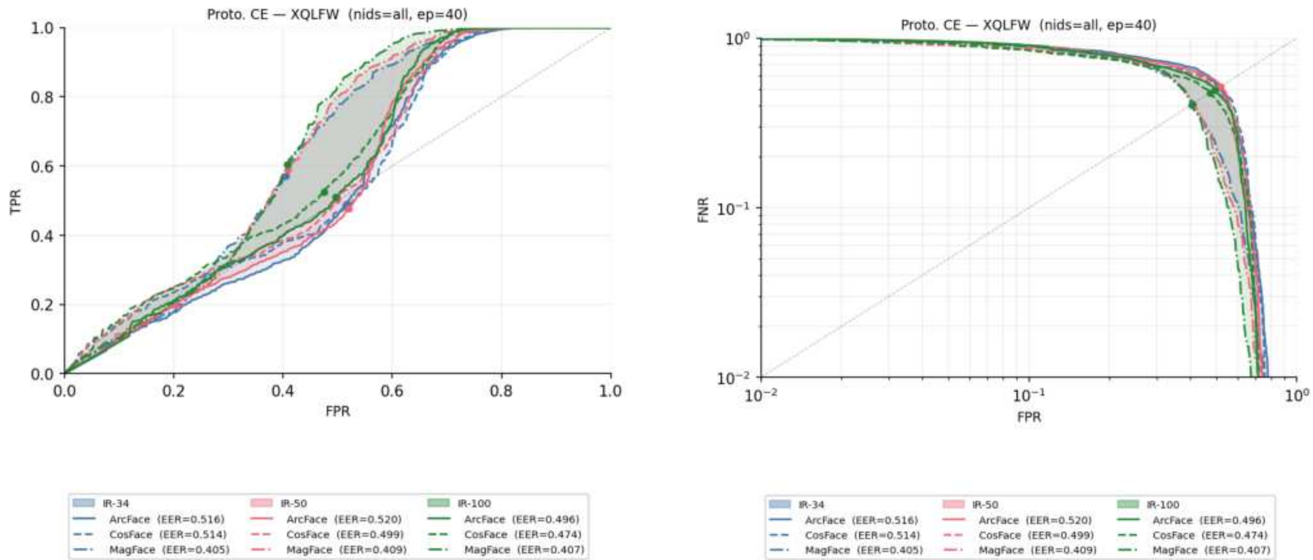


Figure 53. **Threshold-classifier ROC and DET — ProtoCE, XQLFW.** Each curve represents one model configuration (backbone $\times$ loss head $\times$ n_{ids} $\times$ epoch). ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

2.2. Grouped by backbone and loss head

2.2.1 PairCos ($\bar{s}$)

IResNet-34 / ArcFace

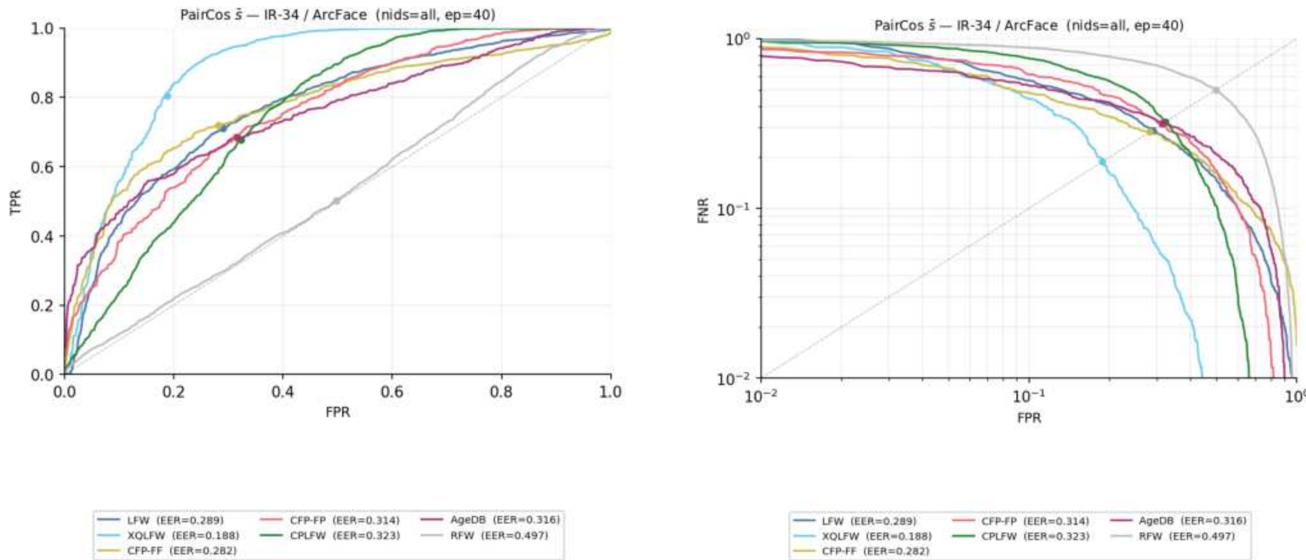


Figure 54. **Threshold-classifier ROC and DET — PairCos ($\bar{s}$), IResNet-34 / ArcFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-34 / CosFace

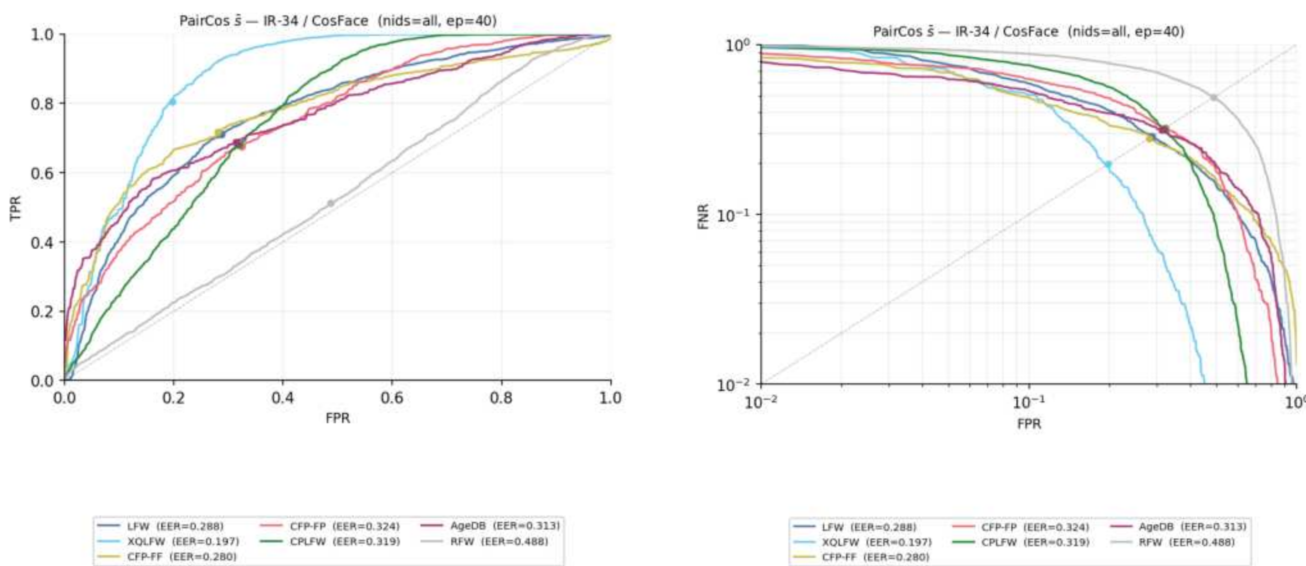


Figure 55. **Threshold-classifier ROC and DET — PairCos ($\bar{s}$), IResNet-34 / CosFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-34 / MagFace

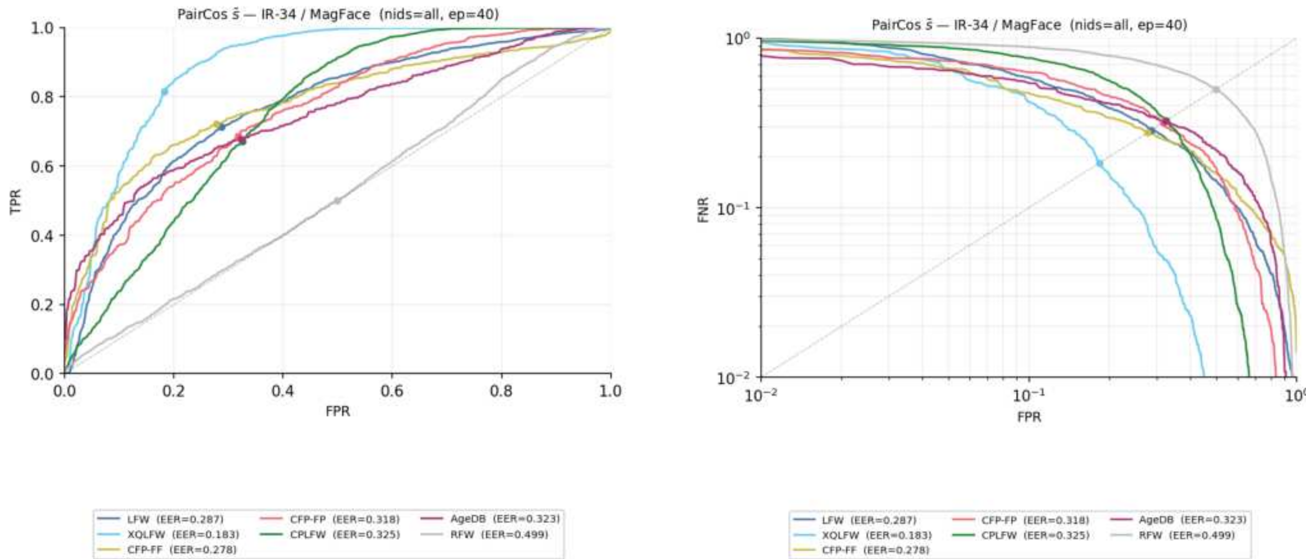


Figure 56. **Threshold-classifier ROC and DET — PairCos ($\bar{s}$), IResNet-34 / MagFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-50 / ArcFace

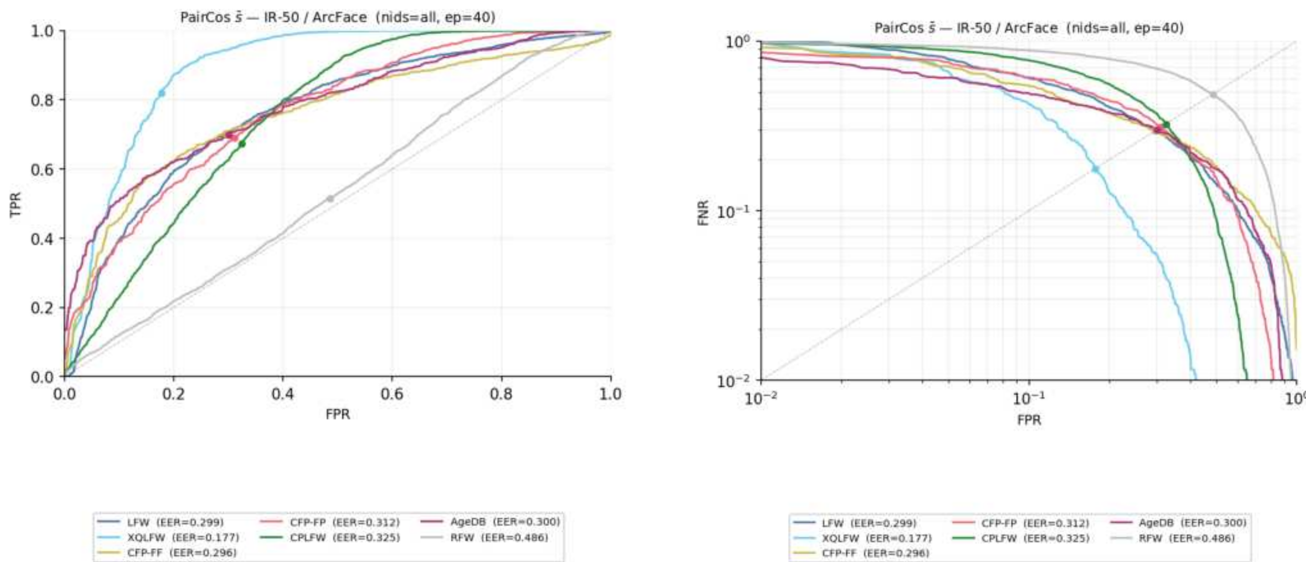


Figure 57. **Threshold-classifier ROC and DET — PairCos ($\bar{s}$), IResNet-50 / ArcFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-50 / CosFace

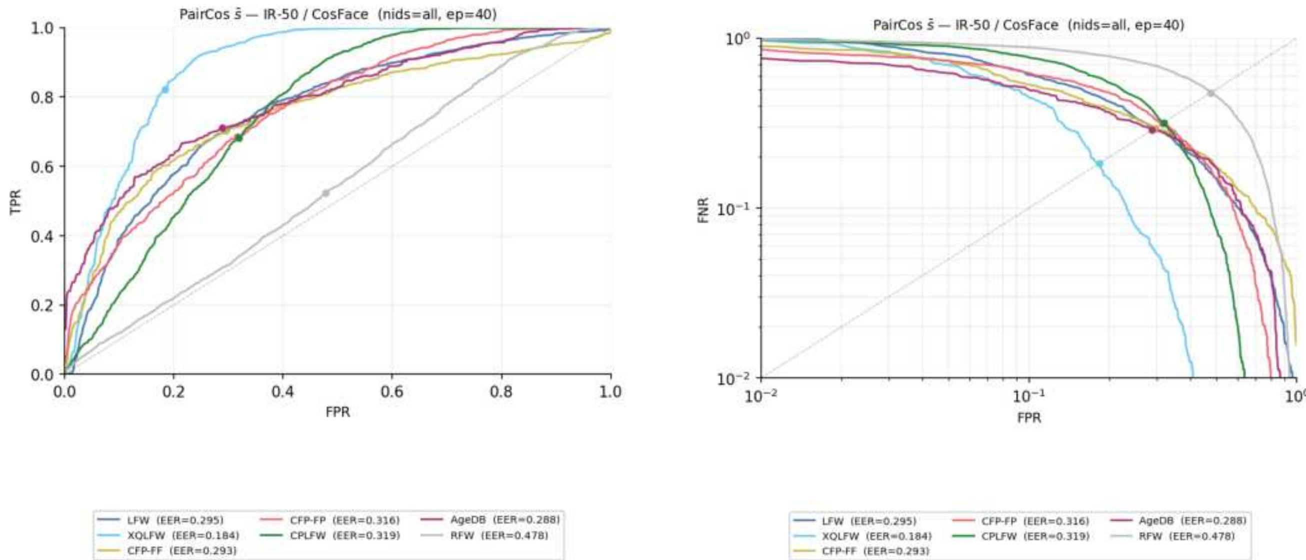


Figure 58. **Threshold-classifier ROC and DET — PairCos ($\bar{s}$), IResNet-50 / CosFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-50 / MagFace

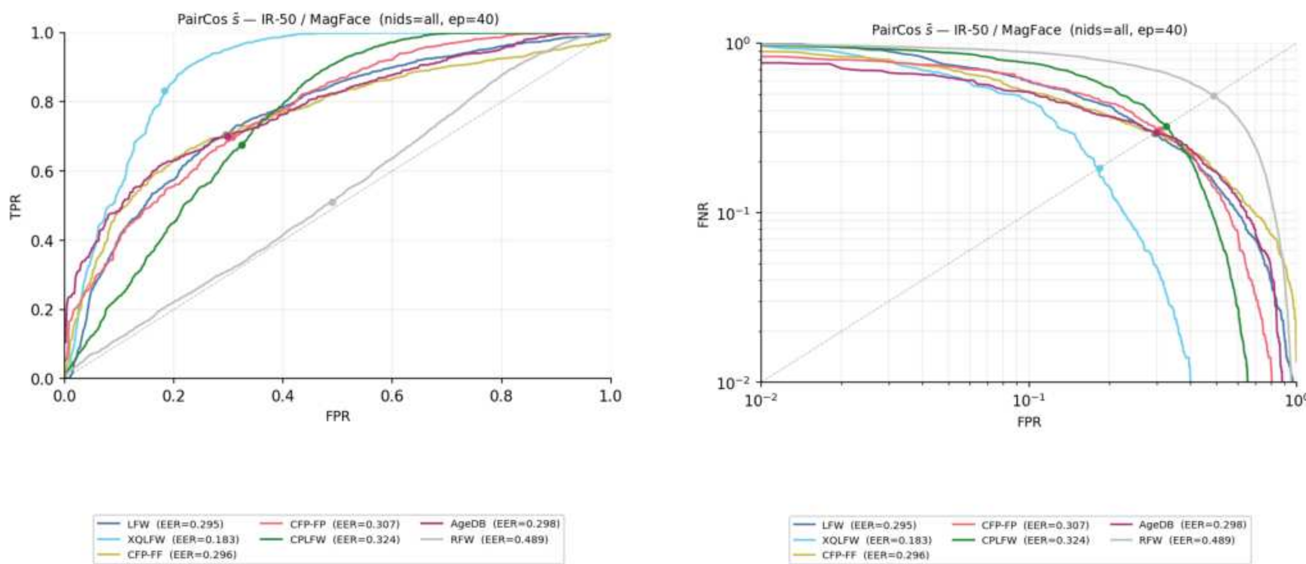


Figure 59. **Threshold-classifier ROC and DET — PairCos ($\bar{s}$), IResNet-50 / MagFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-100 / ArcFace

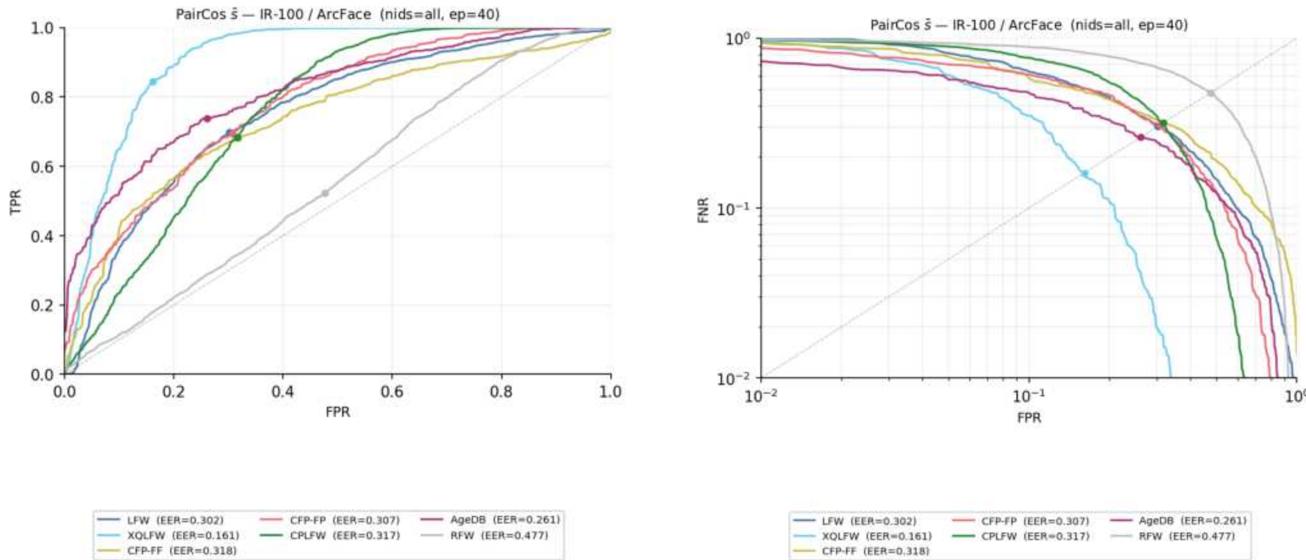


Figure 60. **Threshold-classifier ROC and DET — PairCos ($\bar{s}$), IResNet-100 / ArcFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-100 / CosFace

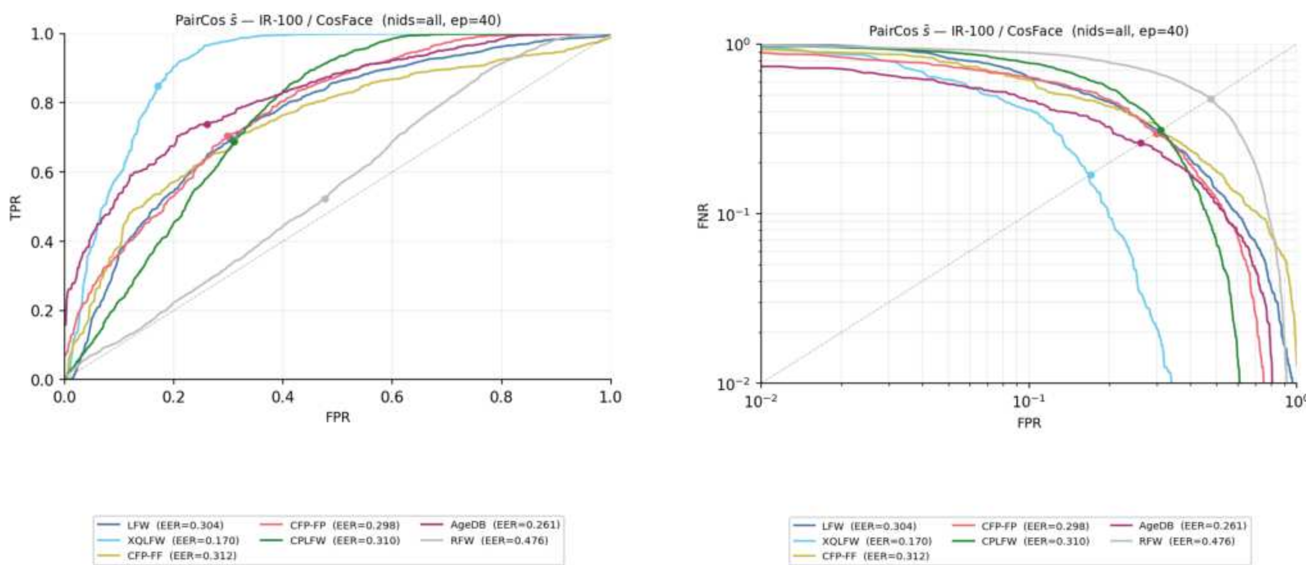


Figure 61. **Threshold-classifier ROC and DET — PairCos ($\bar{s}$), IResNet-100 / CosFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-100 / MagFace

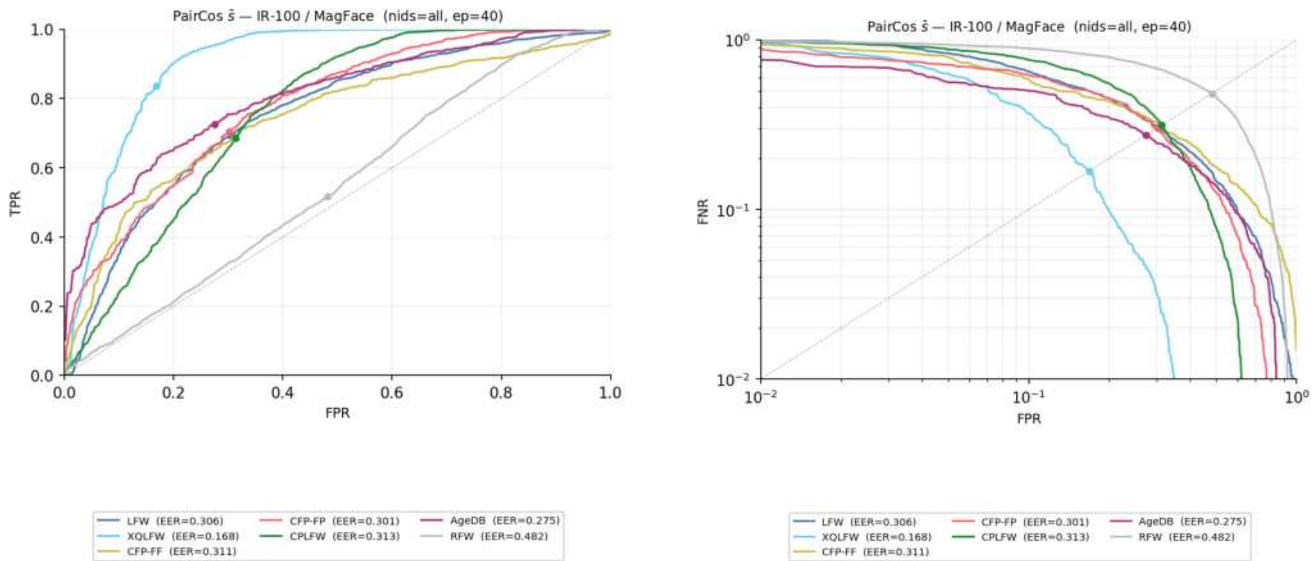


Figure 62. **Threshold-classifier ROC and DET — PairCos ($\bar{s}$), IResNet-100 / MagFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

2.2.2 vMF-κ

IResNet-34 / ArcFace

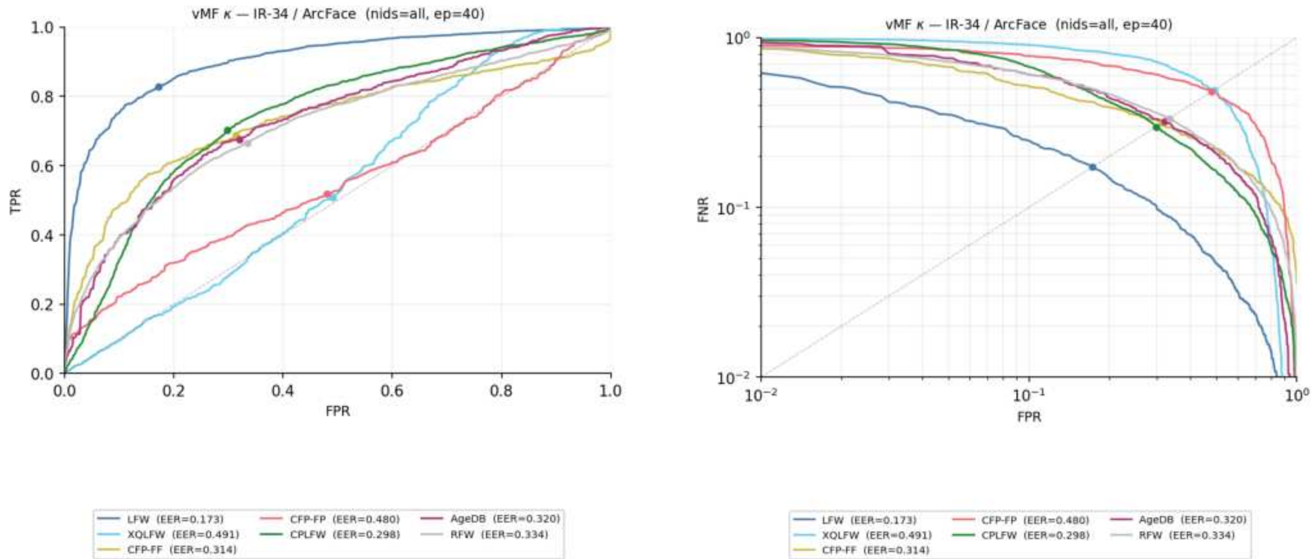


Figure 63. **Threshold-classifier ROC and DET — vMF-κ, IResNet-34 / ArcFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-34 / CosFace

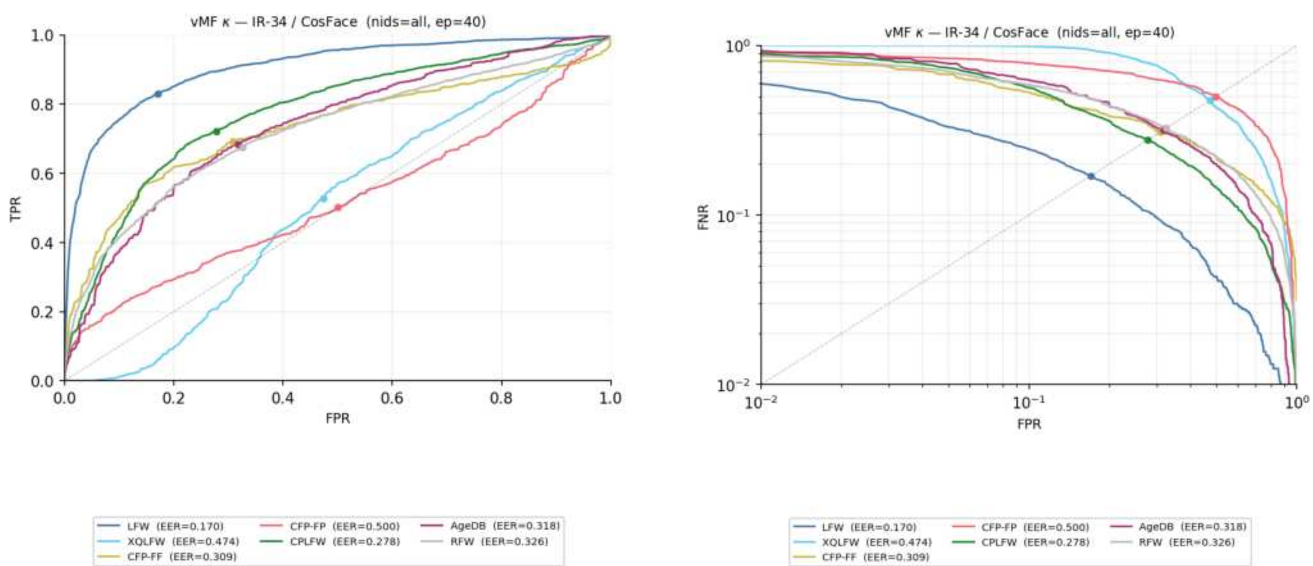


Figure 64. **Threshold-classifier ROC and DET — vMF-κ, IResNet-34 / CosFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-34 / MagFace

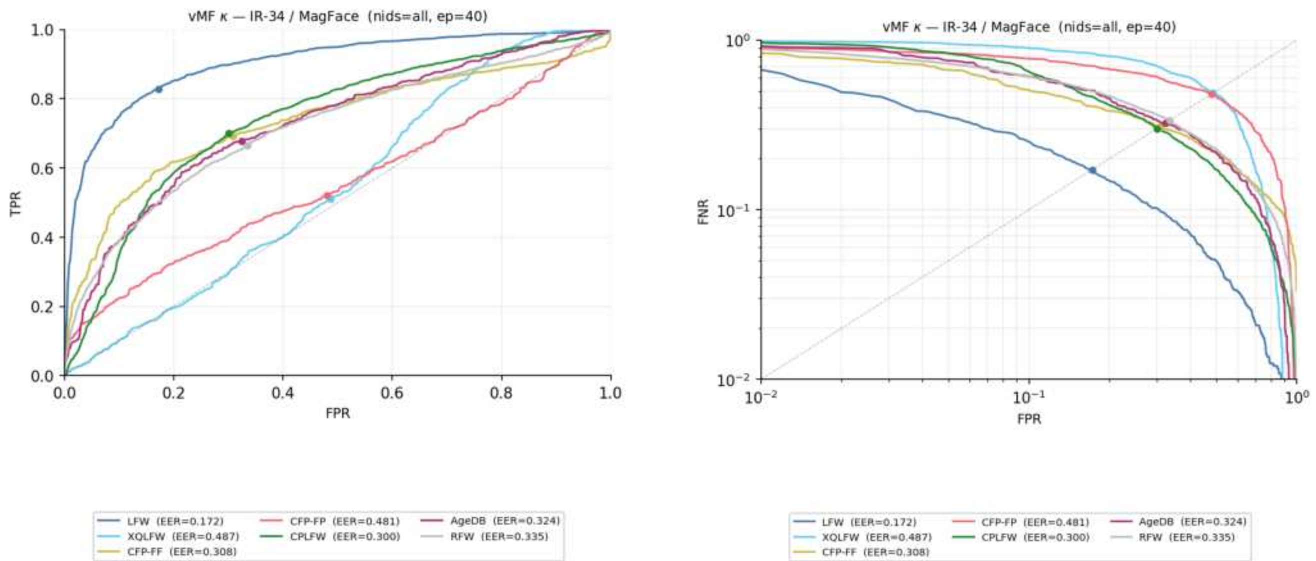


Figure 65. **Threshold-classifier ROC and DET — vMF-κ, IResNet-34 / MagFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-50 / ArcFace

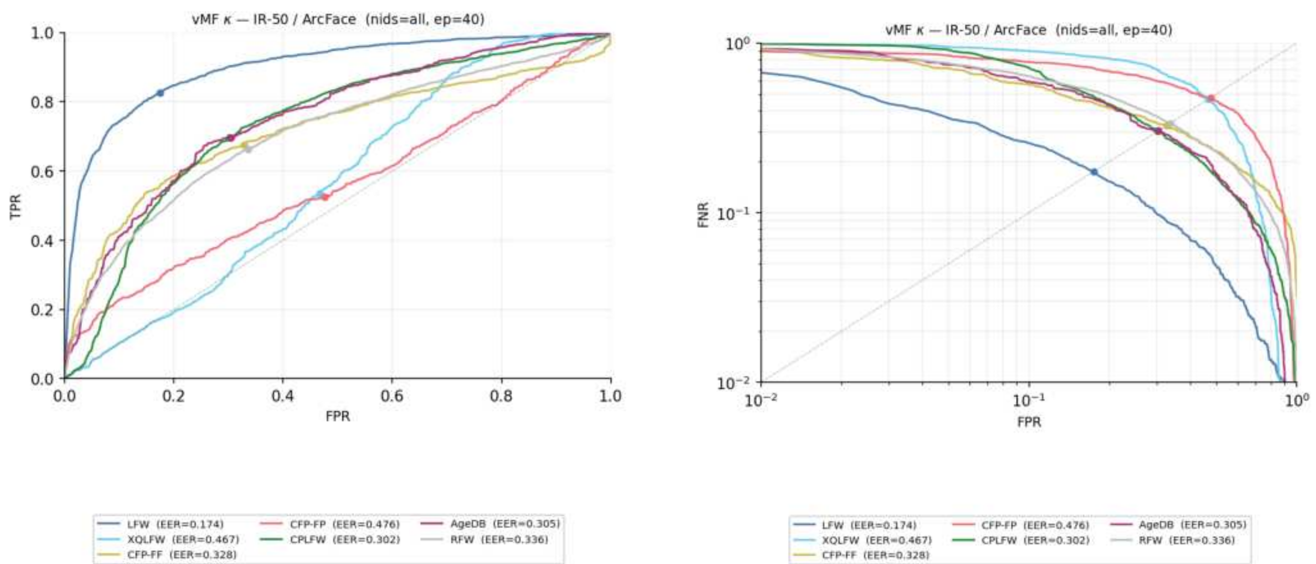


Figure 66. **Threshold-classifier ROC and DET — vMF-κ, IResNet-50 / ArcFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-50 / CosFace

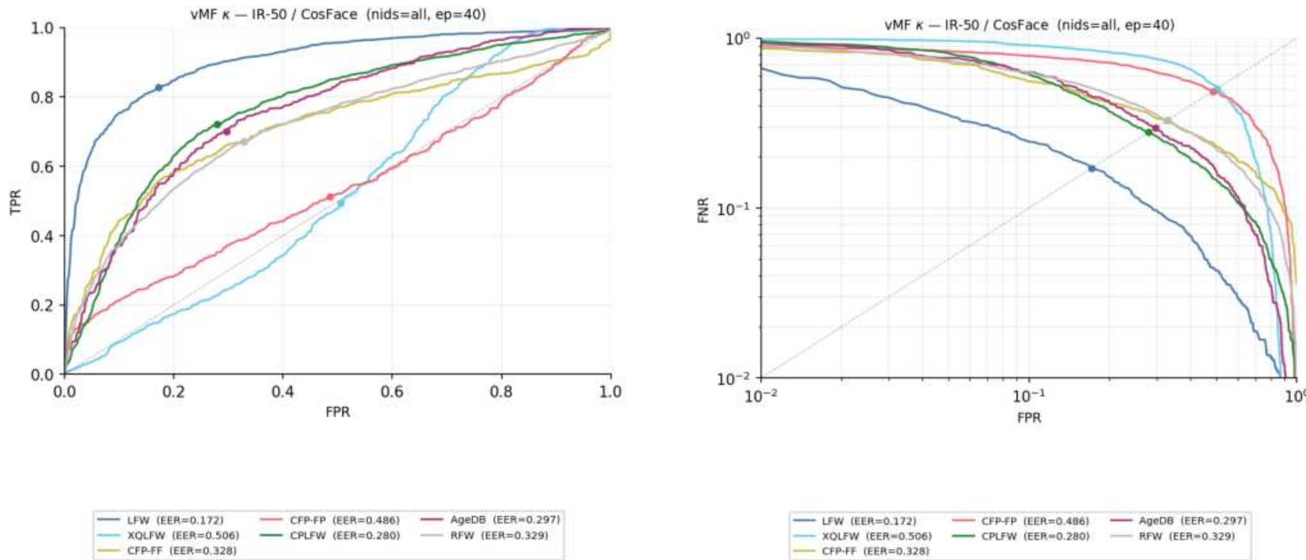


Figure 67. **Threshold-classifier ROC and DET — vMF-κ, IResNet-50 / CosFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-50 / MagFace

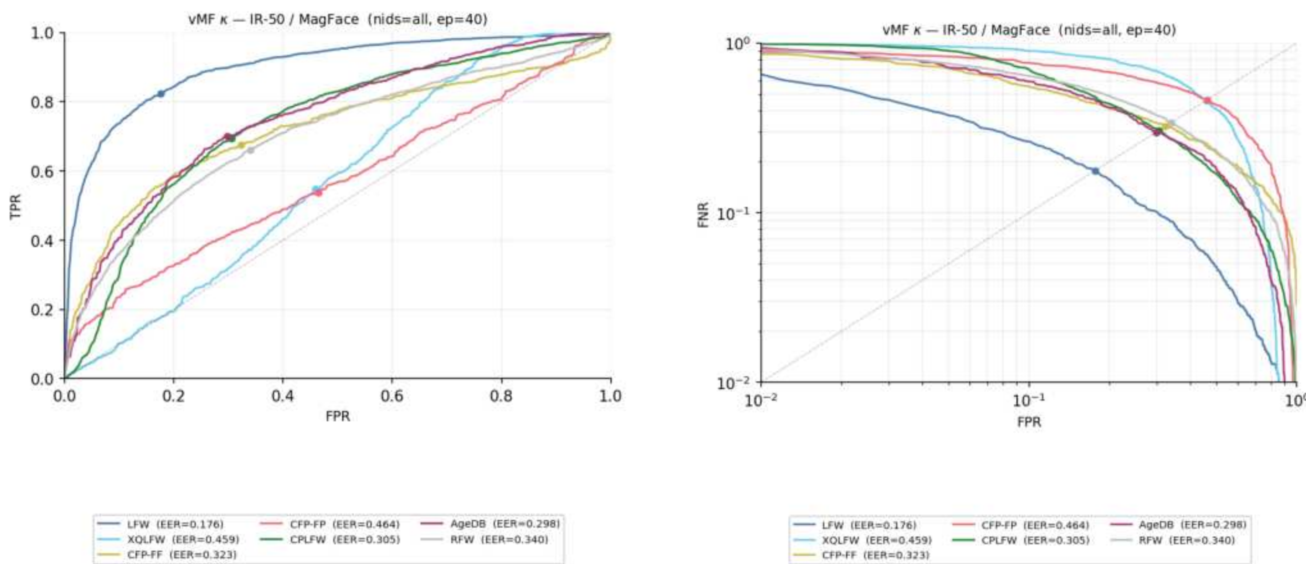


Figure 68. **Threshold-classifier ROC and DET — vMF-κ, IResNet-50 / MagFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-100 / ArcFace

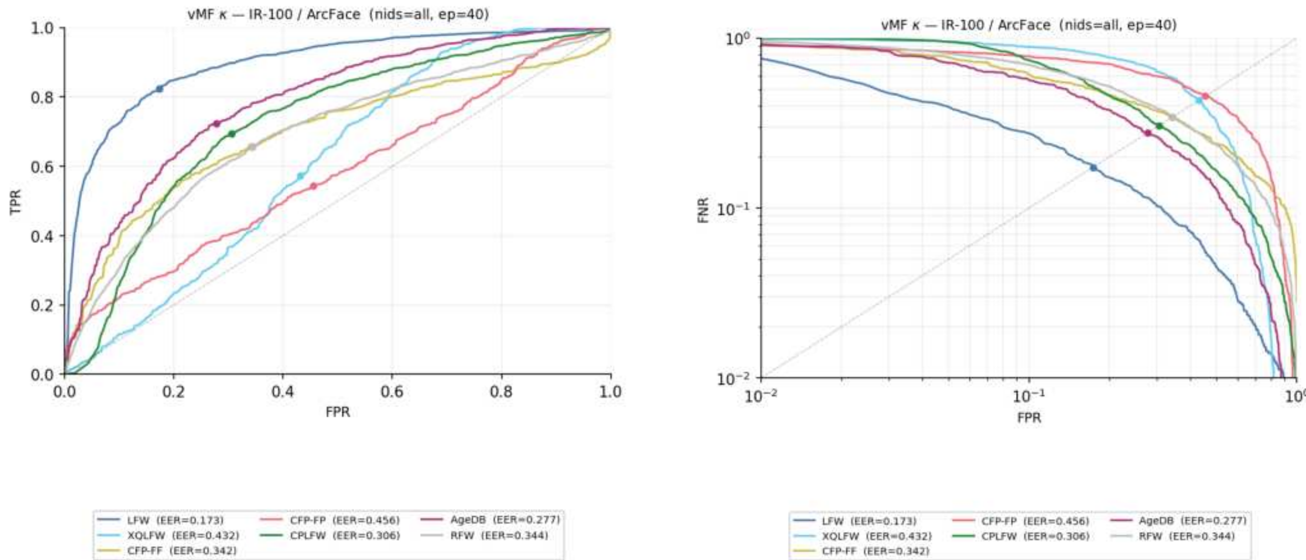


Figure 69. **Threshold-classifier ROC and DET — vMF-κ, IResNet-100 / ArcFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-100 / CosFace

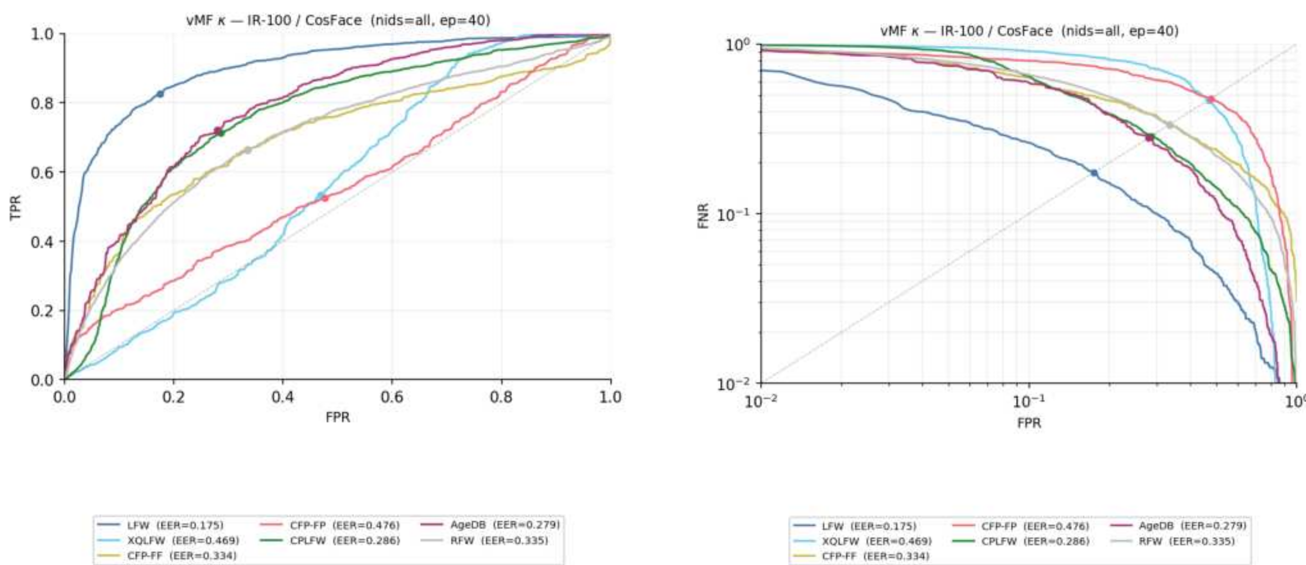


Figure 70. **Threshold-classifier ROC and DET — vMF-κ, IResNet-100 / CosFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-100 / MagFace

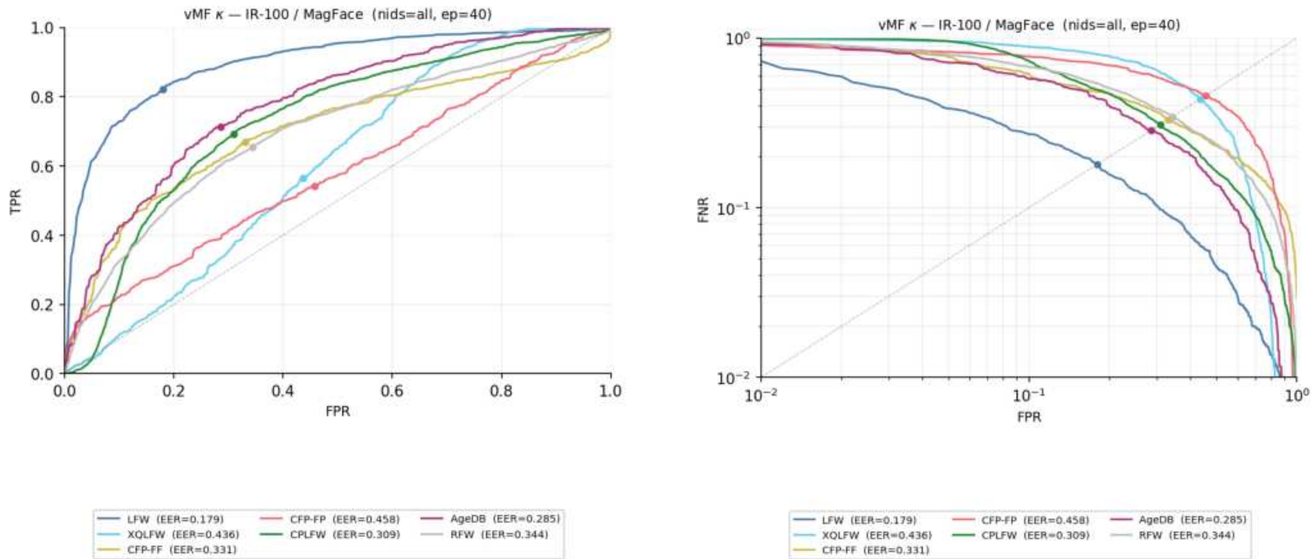


Figure 71. **Threshold-classifier ROC and DET — vMF-κ, IResNet-100 / MagFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

2.2.3 PenLogit

IResNet-34 / ArcFace

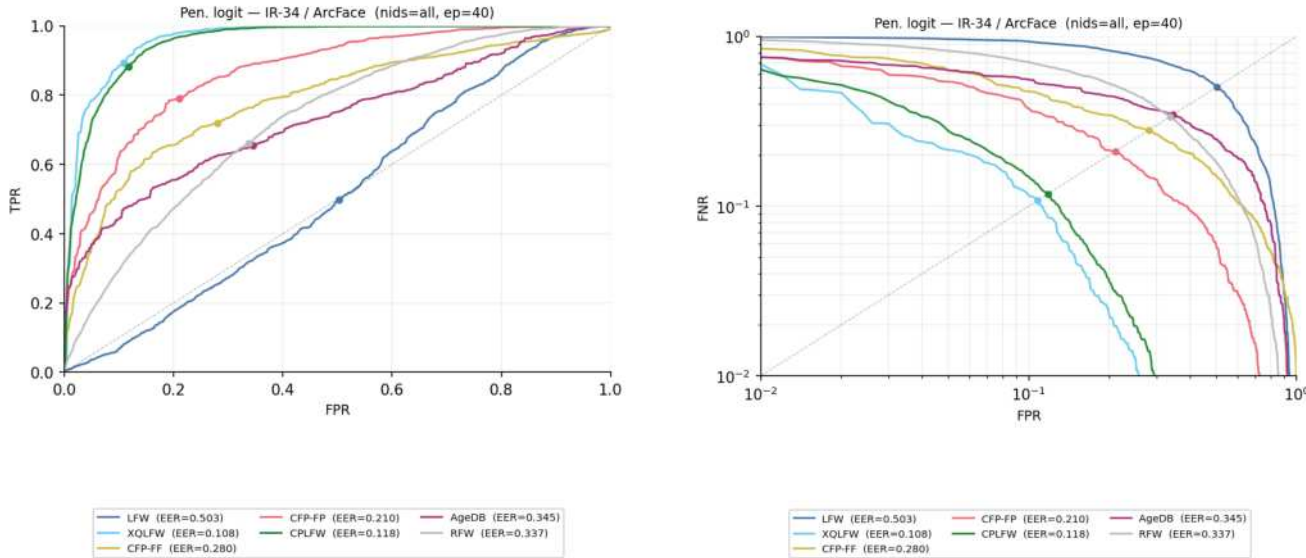


Figure 72. **Threshold-classifier ROC and DET — PenLogit, IResNet-34 / ArcFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-34 / CosFace

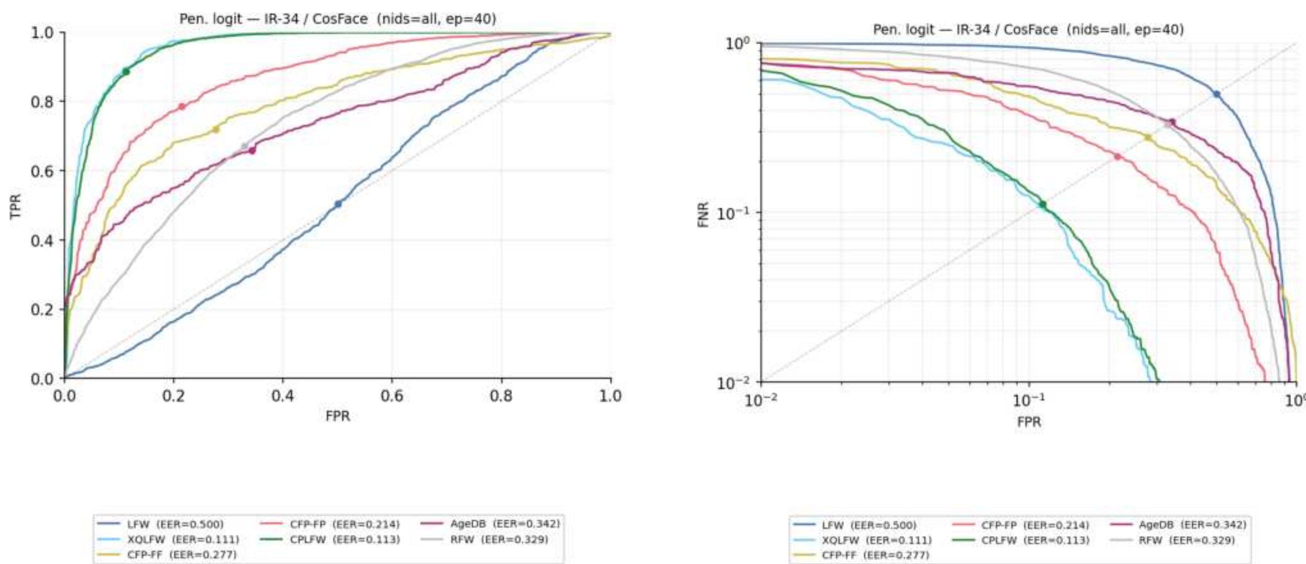


Figure 73. **Threshold-classifier ROC and DET — PenLogit, IResNet-34 / CosFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-34 / MagFace

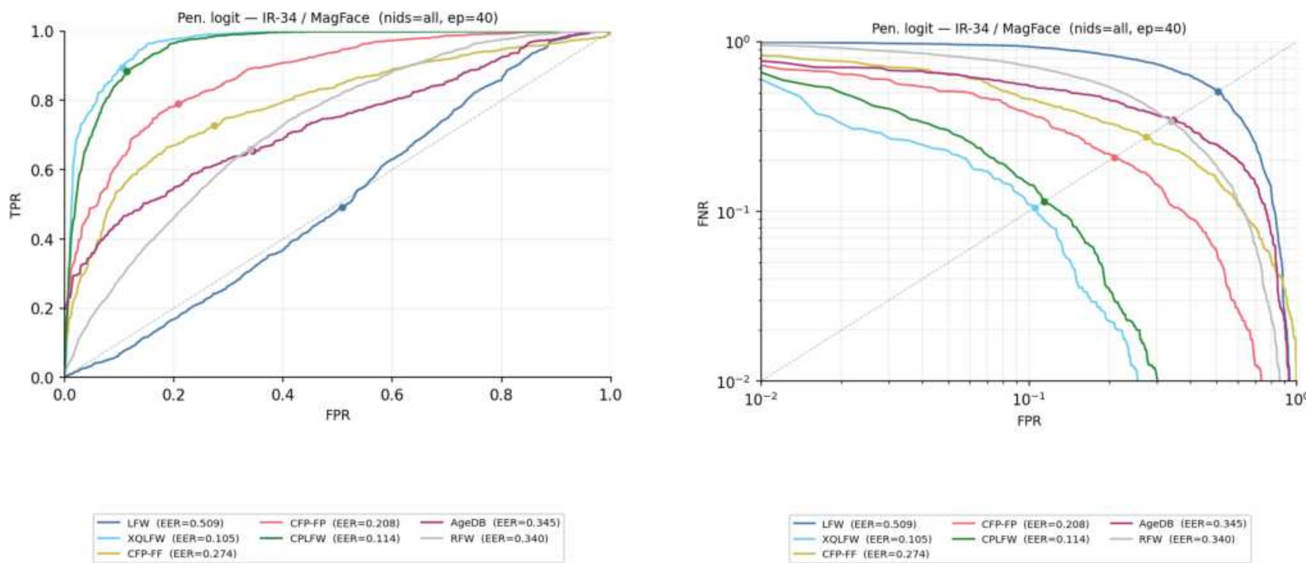


Figure 74. **Threshold-classifier ROC and DET — PenLogit, IResNet-34 / MagFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-50 / ArcFace

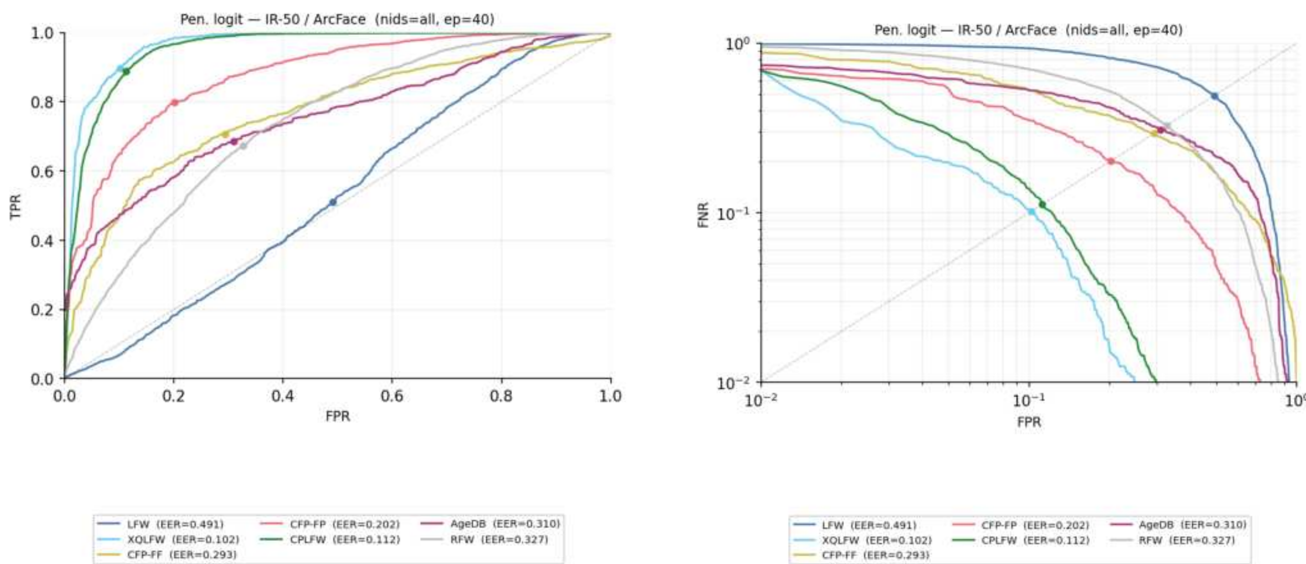


Figure 75. **Threshold-classifier ROC and DET — PenLogit, IResNet-50 / ArcFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-50 / CosFace

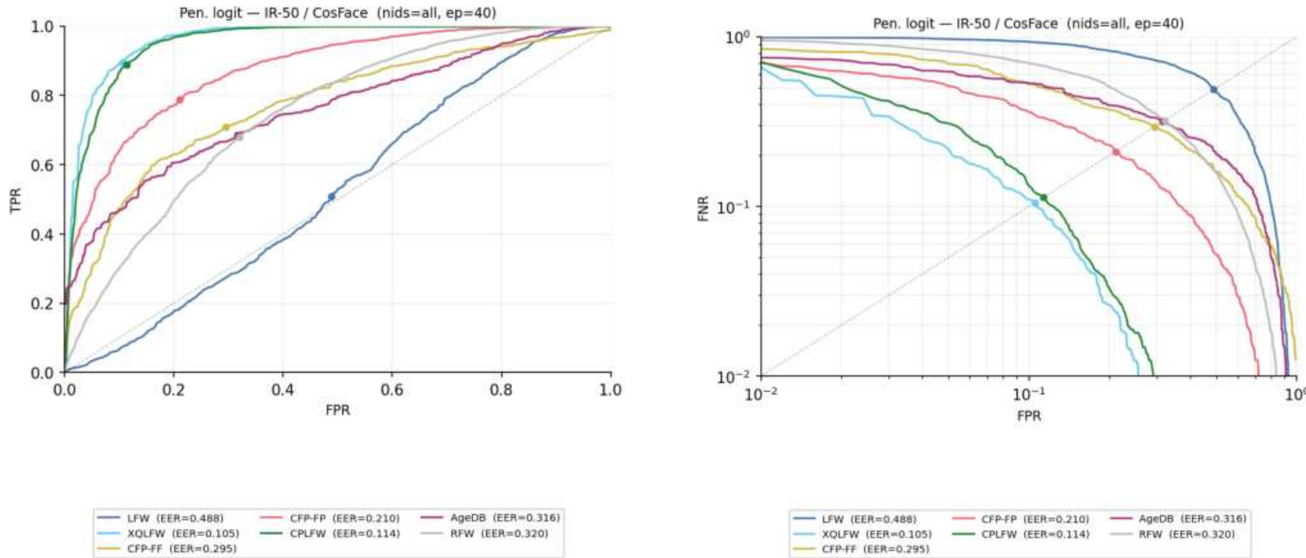


Figure 76. **Threshold-classifier ROC and DET — PenLogit, IResNet-50 / CosFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-50 / MagFace

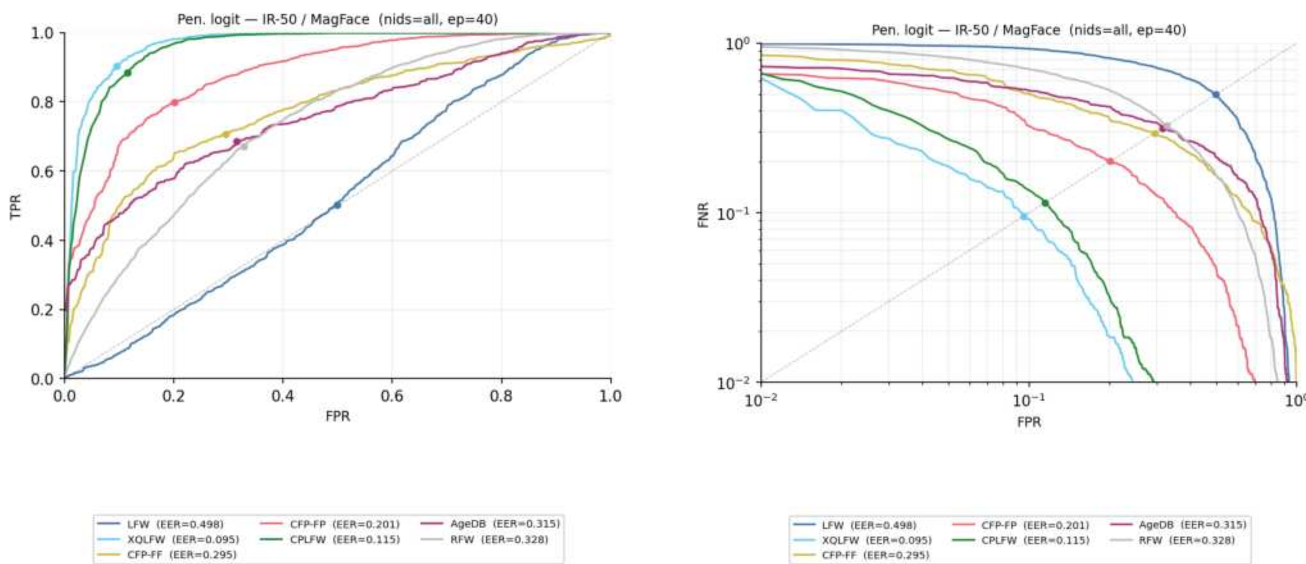


Figure 77. **Threshold-classifier ROC and DET — PenLogit, IResNet-50 / MagFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-100 / ArcFace

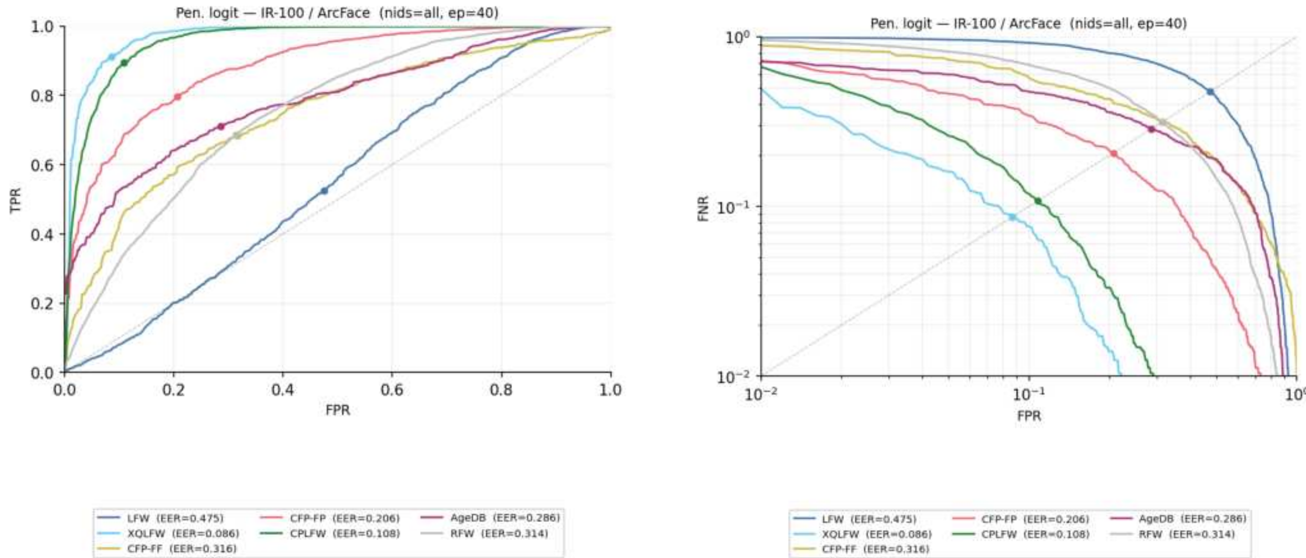


Figure 78. **Threshold-classifier ROC and DET — PenLogit, IResNet-100 / ArcFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-100 / CosFace

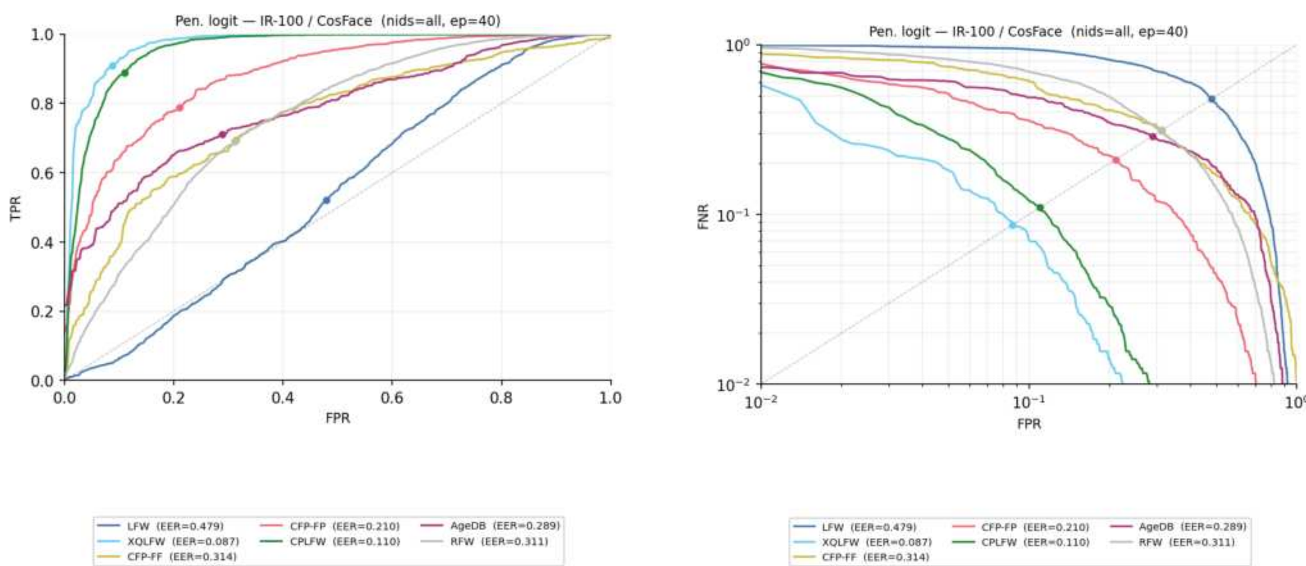


Figure 79. **Threshold-classifier ROC and DET — PenLogit, IResNet-100 / CosFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-100 / MagFace

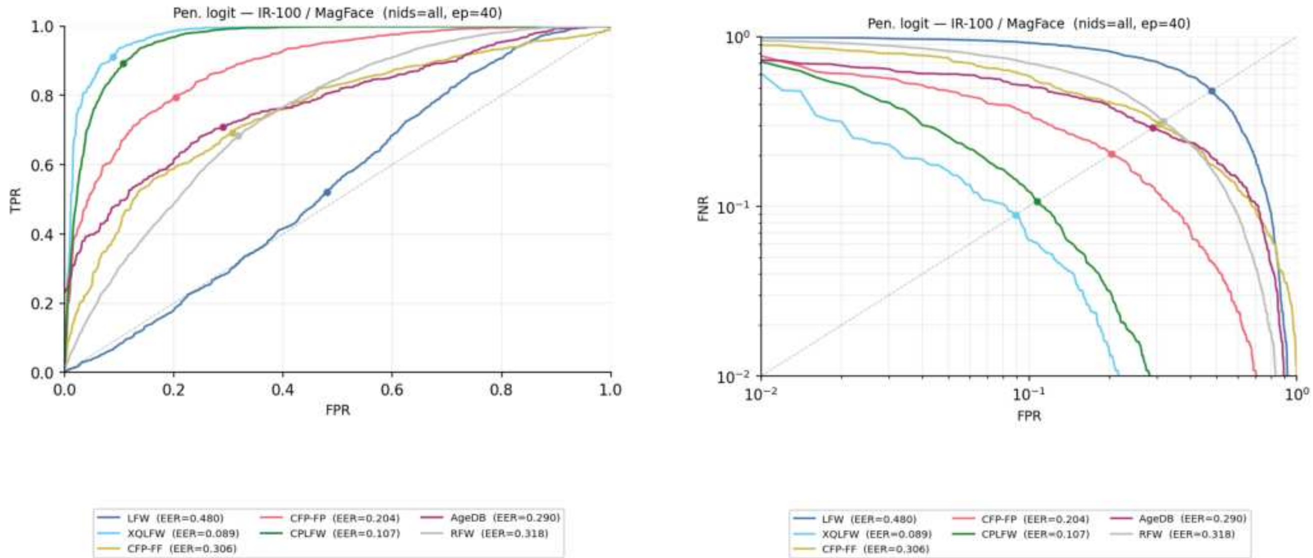


Figure 80. **Threshold-classifier ROC and DET — PenLogit, IResNet-100 / MagFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

2.2.4 ProtoCE

IResNet-34 / ArcFace

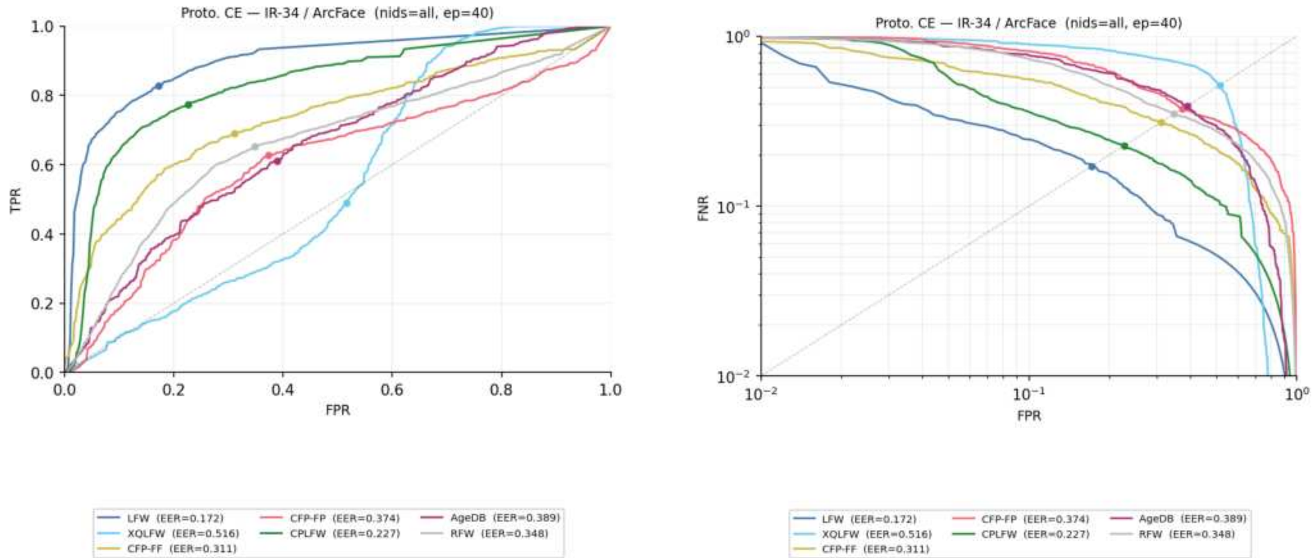


Figure 81. **Threshold-classifier ROC and DET — ProtoCE, IResNet-34 / ArcFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-34 / CosFace

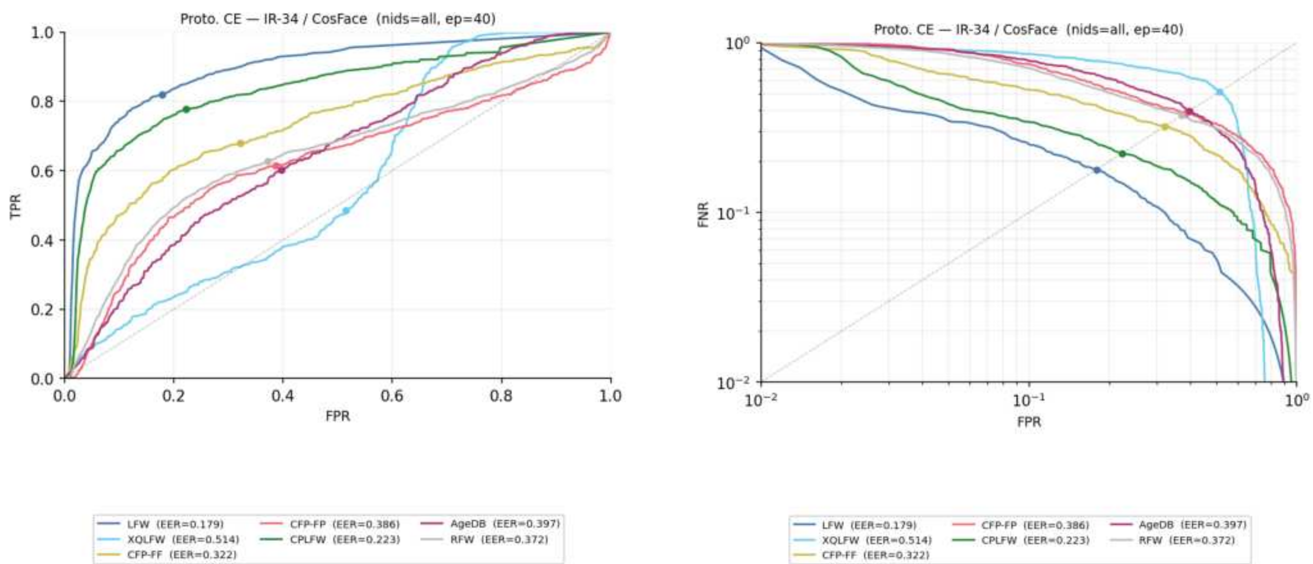


Figure 82. **Threshold-classifier ROC and DET — ProtoCE, IResNet-34 / CosFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-34 / MagFace

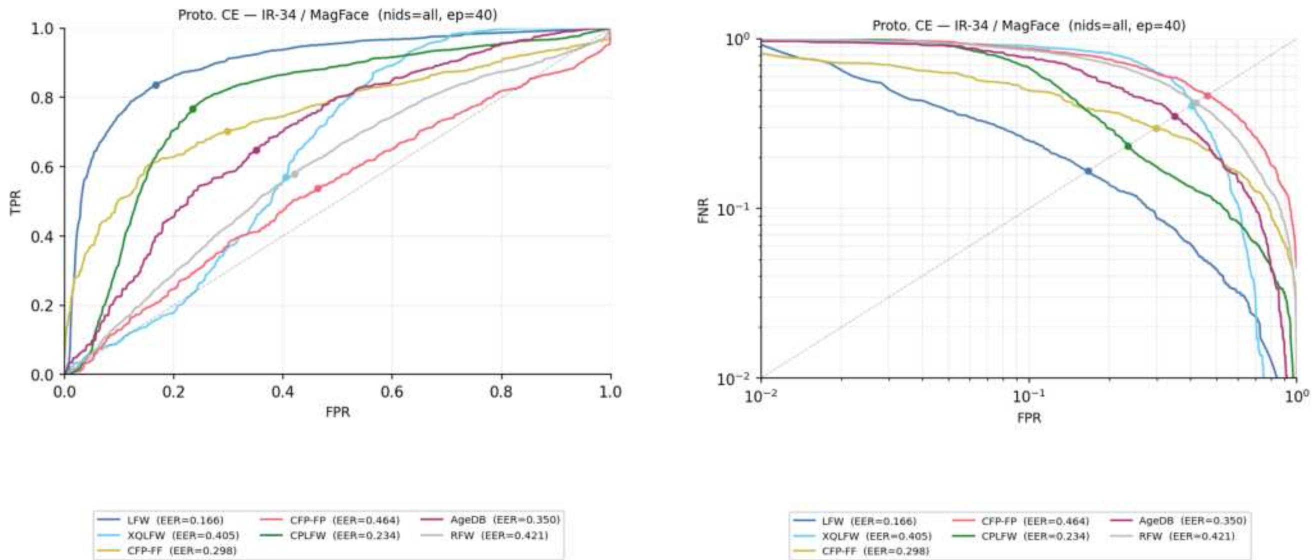


Figure 83. **Threshold-classifier ROC and DET — ProtoCE, IResNet-34 / MagFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-50 / ArcFace

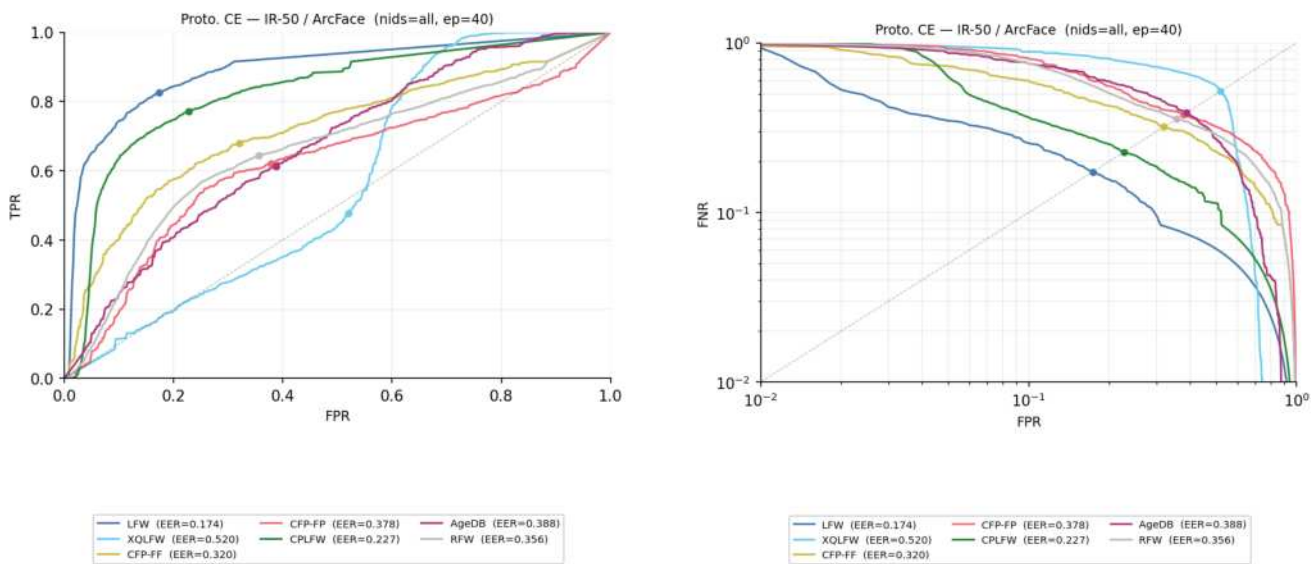


Figure 84. **Threshold-classifier ROC and DET — ProtoCE, IResNet-50 / ArcFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-50 / CosFace

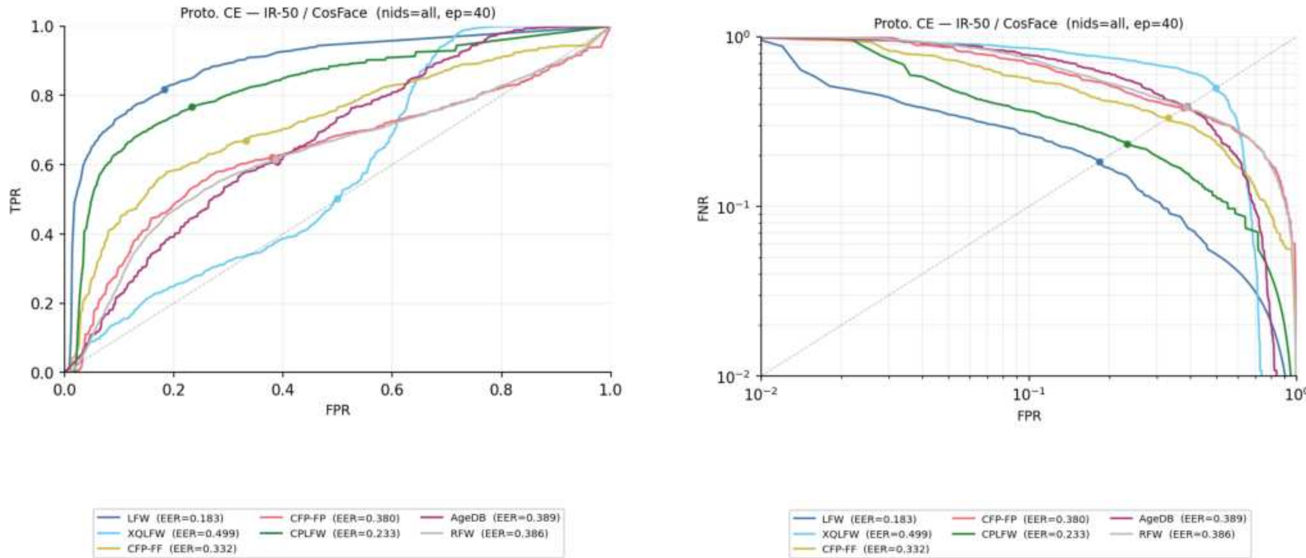


Figure 85. **Threshold-classifier ROC and DET — ProtoCE, IResNet-50 / CosFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-50 / MagFace

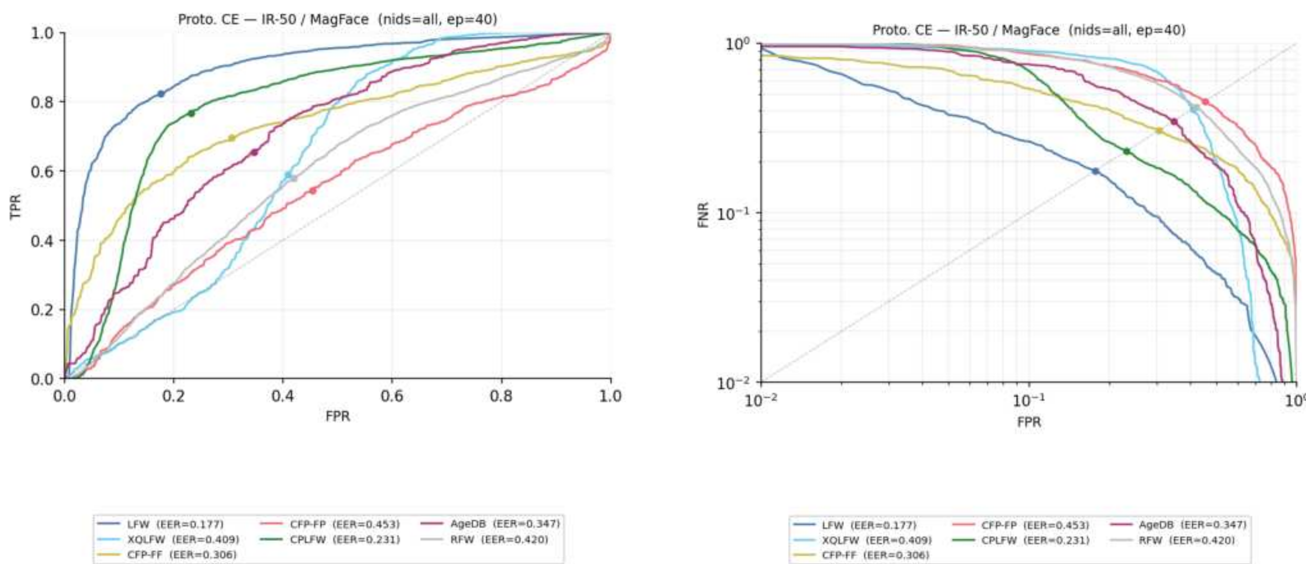


Figure 86. **Threshold-classifier ROC and DET — ProtoCE, IResNet-50 / MagFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-100 / ArcFace

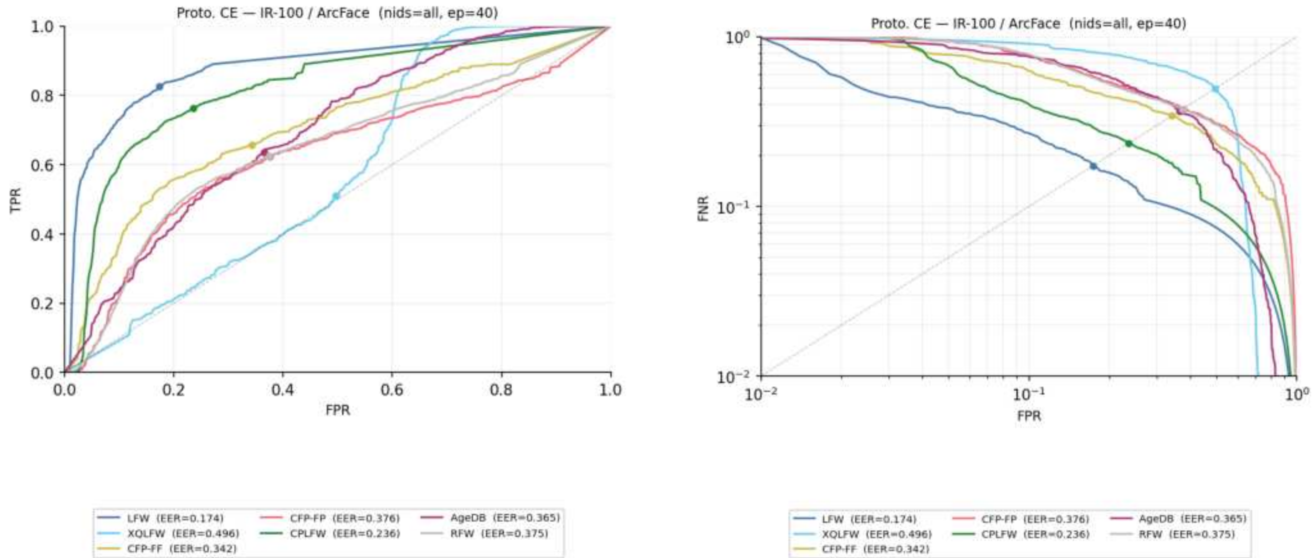


Figure 87. **Threshold-classifier ROC and DET — ProtoCE, IResNet-100 / ArcFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-100 / CosFace

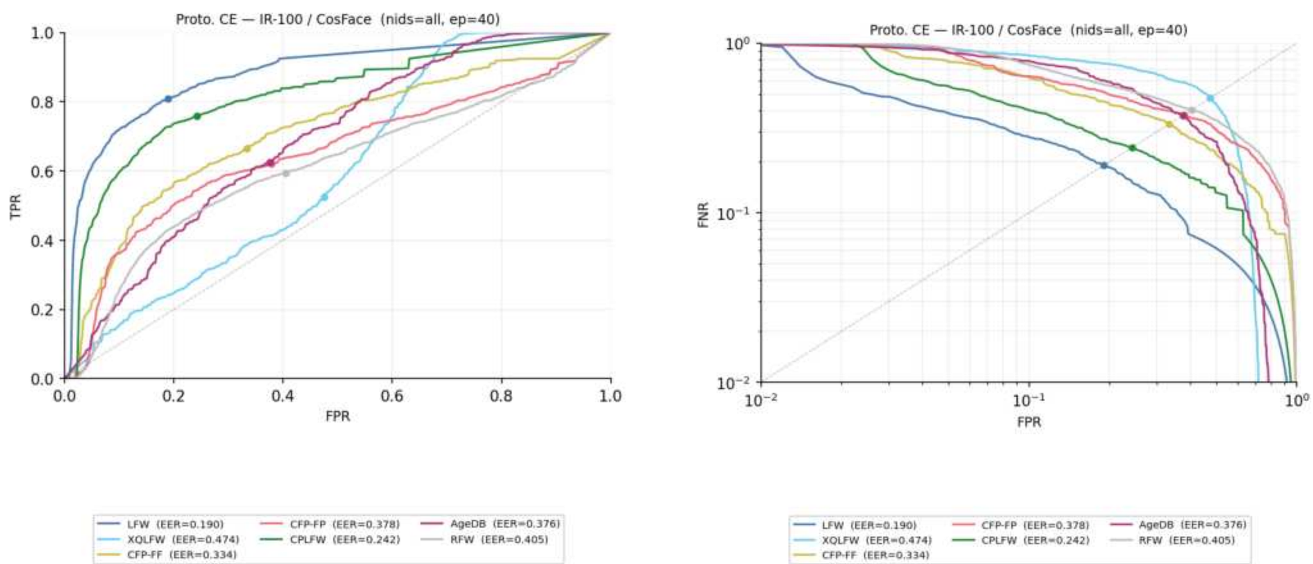


Figure 88. **Threshold-classifier ROC and DET — ProtoCE, IResNet-100 / CosFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

IResNet-100 / MagFace

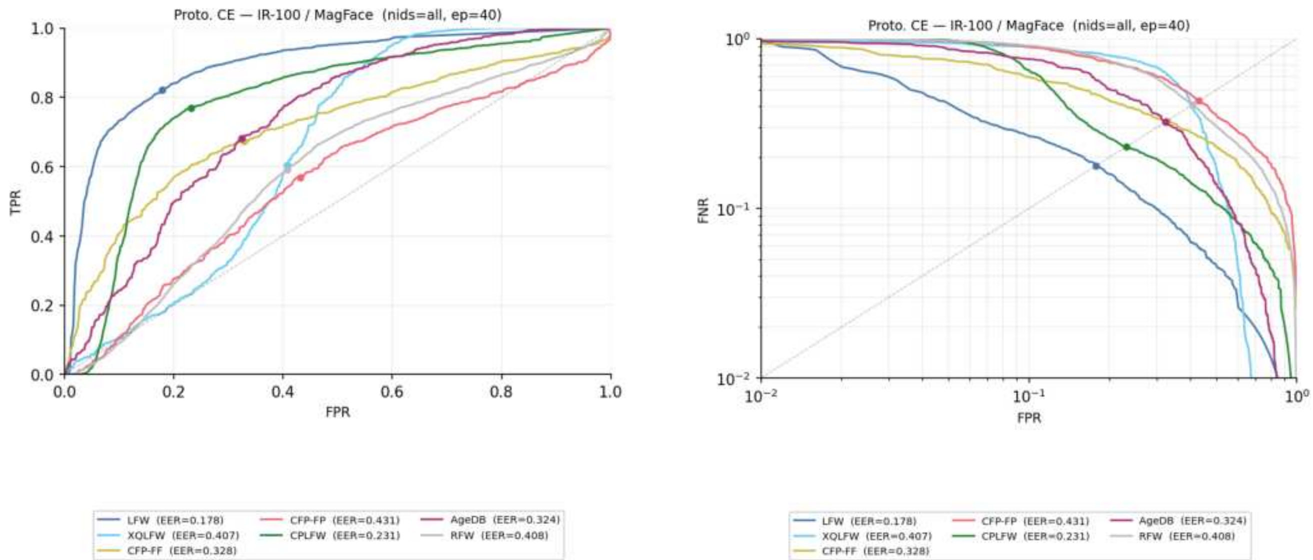


Figure 89. **Threshold-classifier ROC and DET — ProtoCE, IResNet-100 / MagFace.** Each curve corresponds to one combination of training-set size (n_{ids}) and epoch across all benchmarks. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

2.3. Grouped by training-set size

2.3.1 PairCos ($\bar{s}$)

AgeDB-30

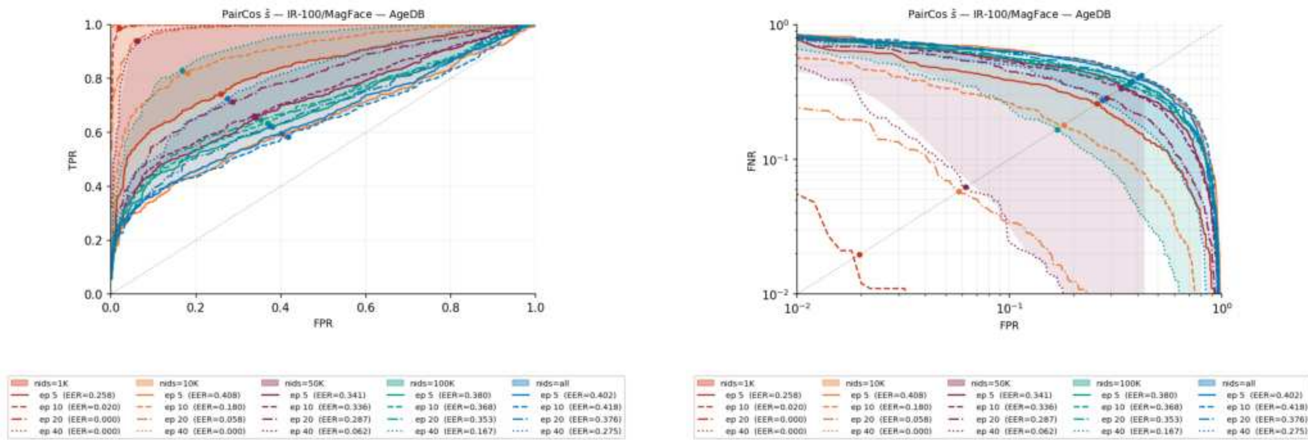


Figure 90. **Threshold-classifier ROC and DET — PairCos ($\bar{s}$), AgeDB-30.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

CFP-FF

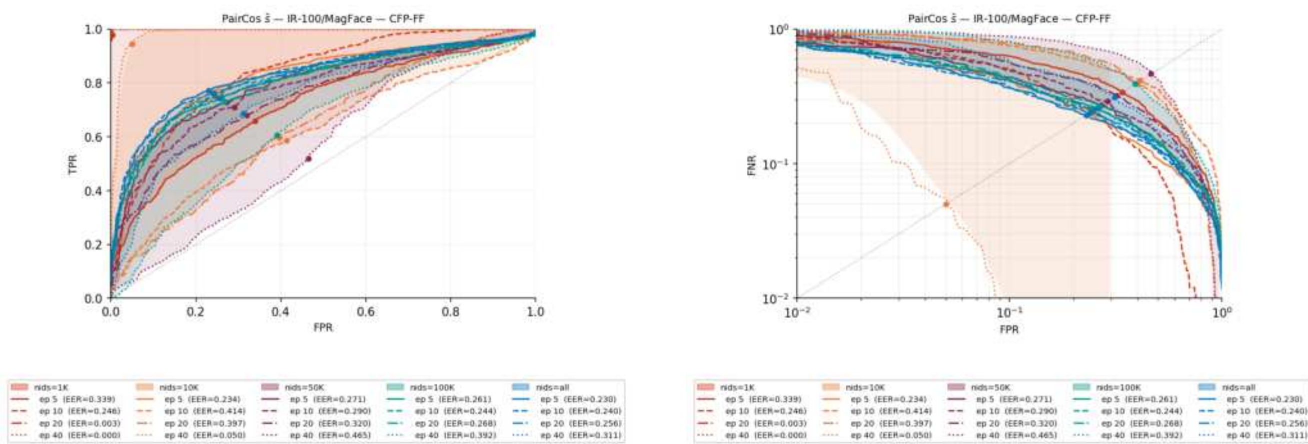


Figure 91. **Threshold-classifier ROC and DET — PairCos ($\bar{s}$), CFP-FF.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

CFP-FP

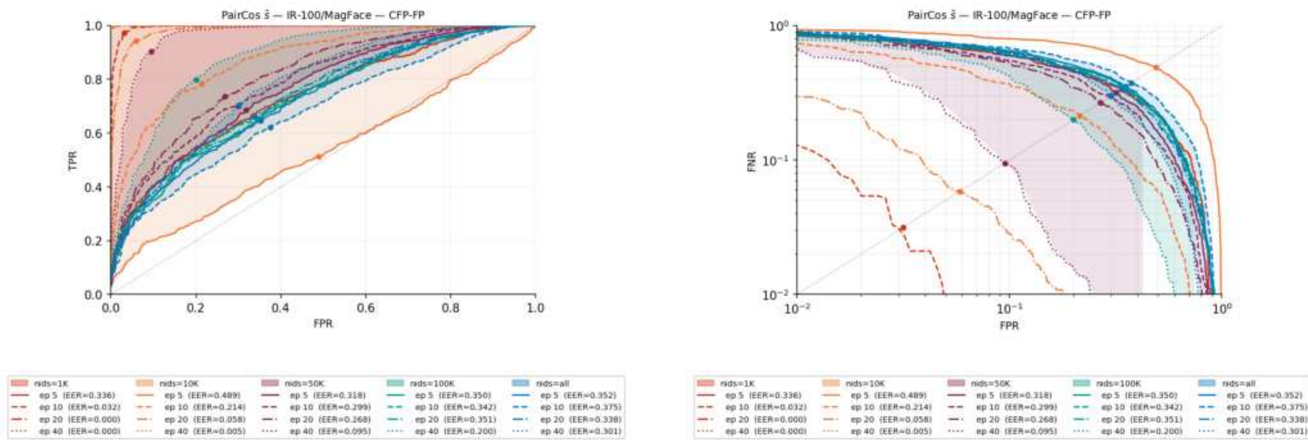


Figure 92. **Threshold-classifier ROC and DET — PairCos ($\bar{s}$), CFP-FP.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

CPLFW

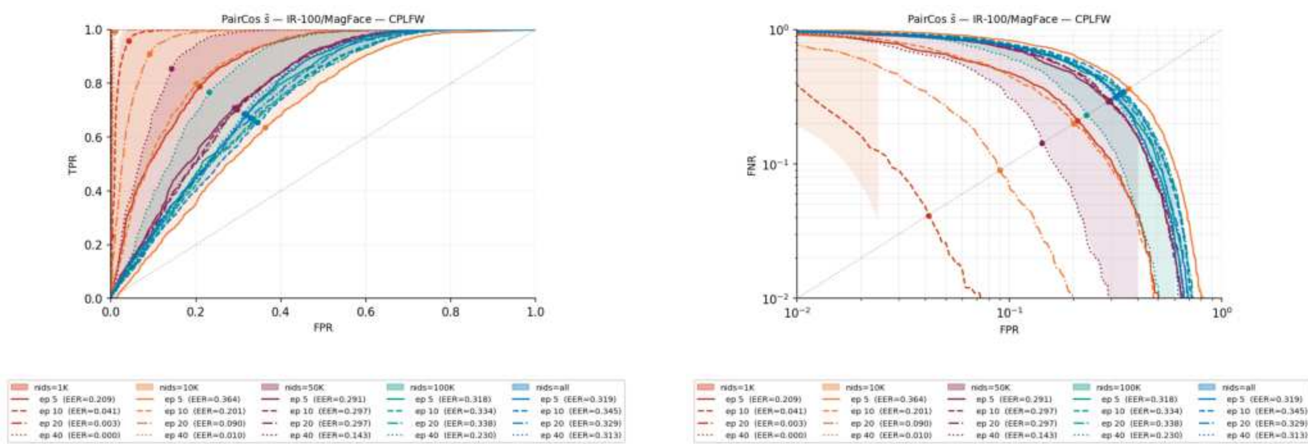


Figure 93. **Threshold-classifier ROC and DET — PairCos ($\bar{s}$), CPLFW.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

LFW

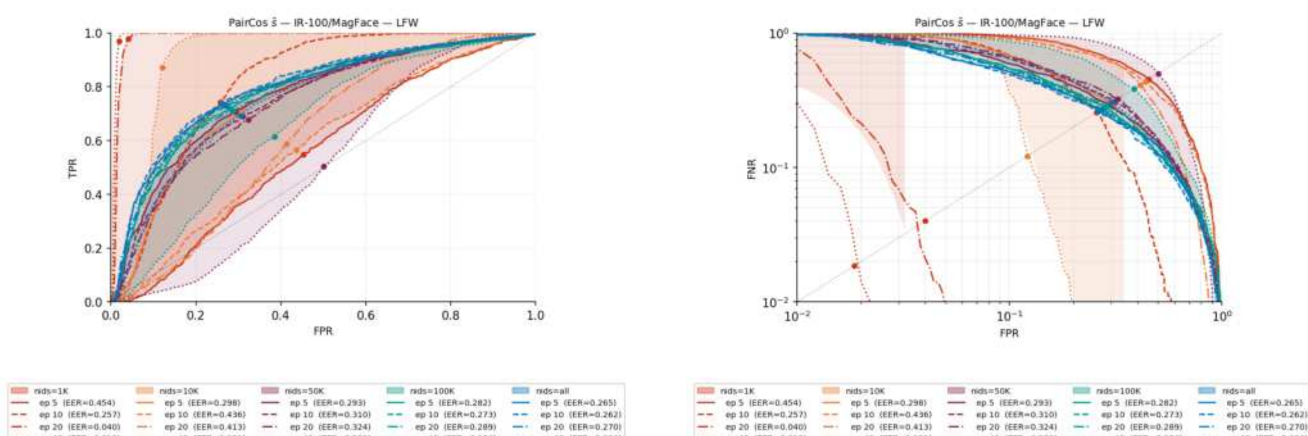


Figure 94. **Threshold-classifier ROC and DET — PairCos ($\bar{s}$), LFW.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

RFW

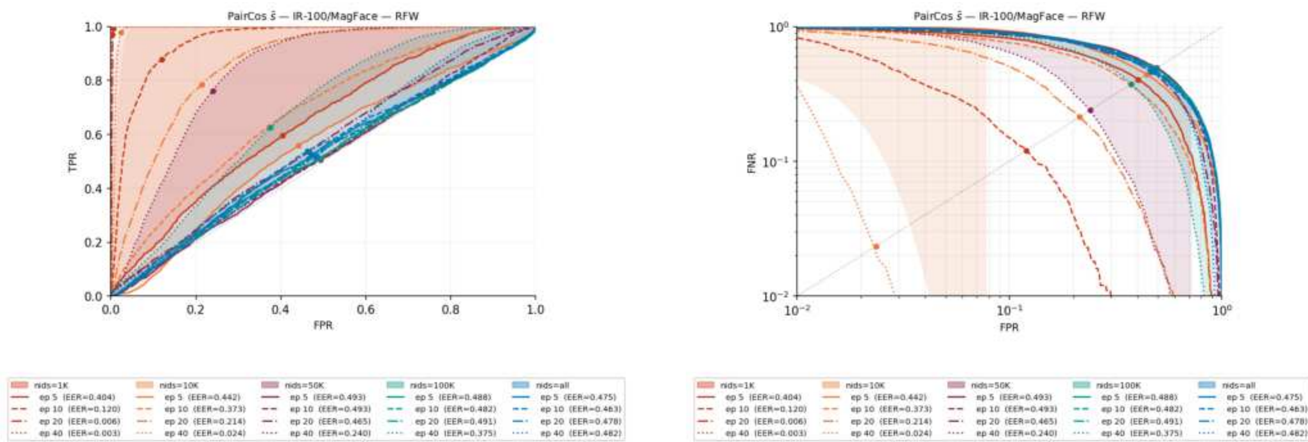


Figure 95. **Threshold-classifier ROC and DET — PairCos ($\bar{s}$), RFW.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K},\ 10\text{K},\ 50\text{K},\ 100\text{K},\ \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

XQLFW

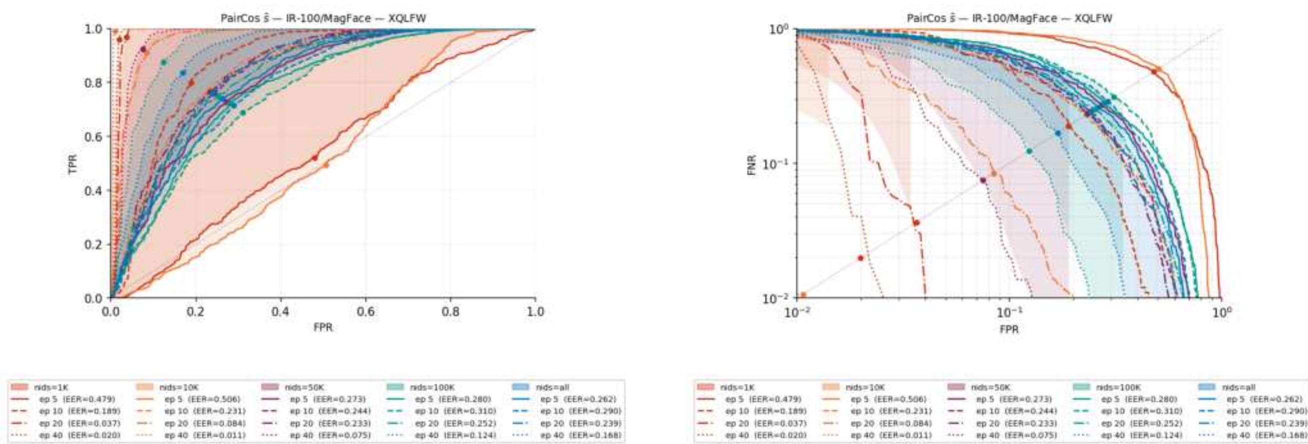


Figure 96. **Threshold-classifier ROC and DET — PairCos ($\bar{s}$), XQLFW.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K},\ 10\text{K},\ 50\text{K},\ 100\text{K},\ \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

WebFace4M (held-out)

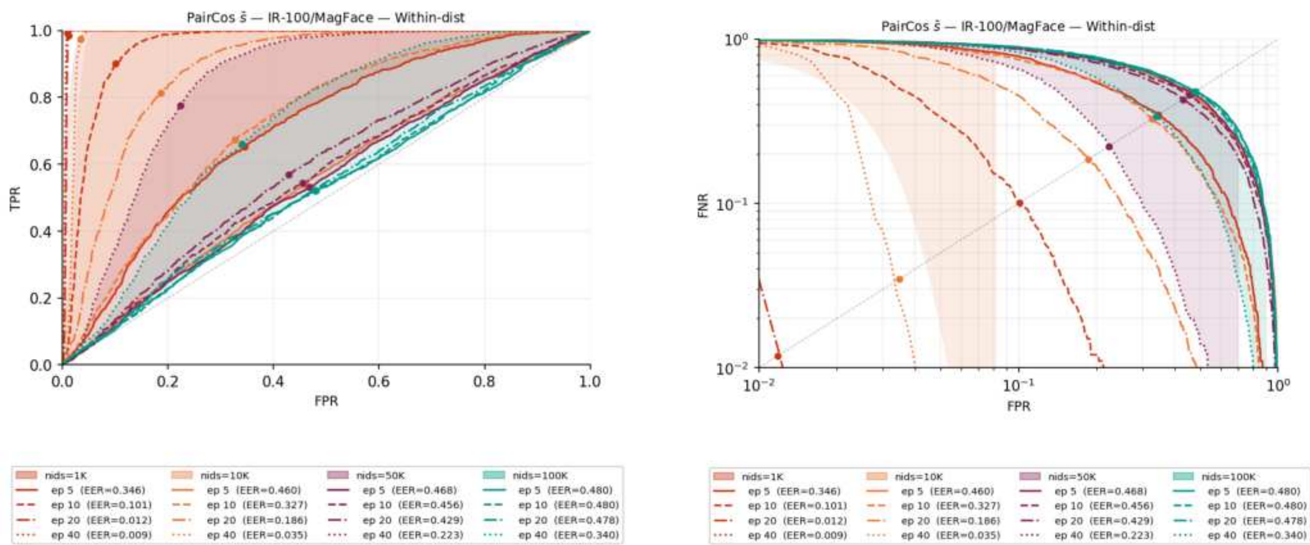


Figure 97. **Threshold-classifier ROC and DET — PairCos ($\bar{s}$), WebFace4M (held-out).** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

2.3.2 vMF-κ

AgeDB-30

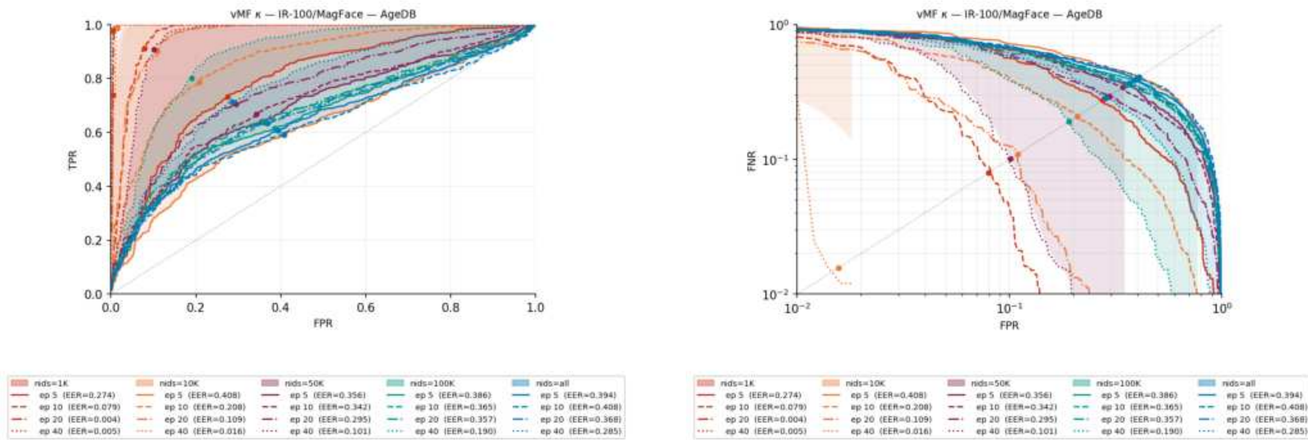


Figure 98. **Threshold-classifier ROC and DET — vMF-κ, AgeDB-30.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

CFP-FF

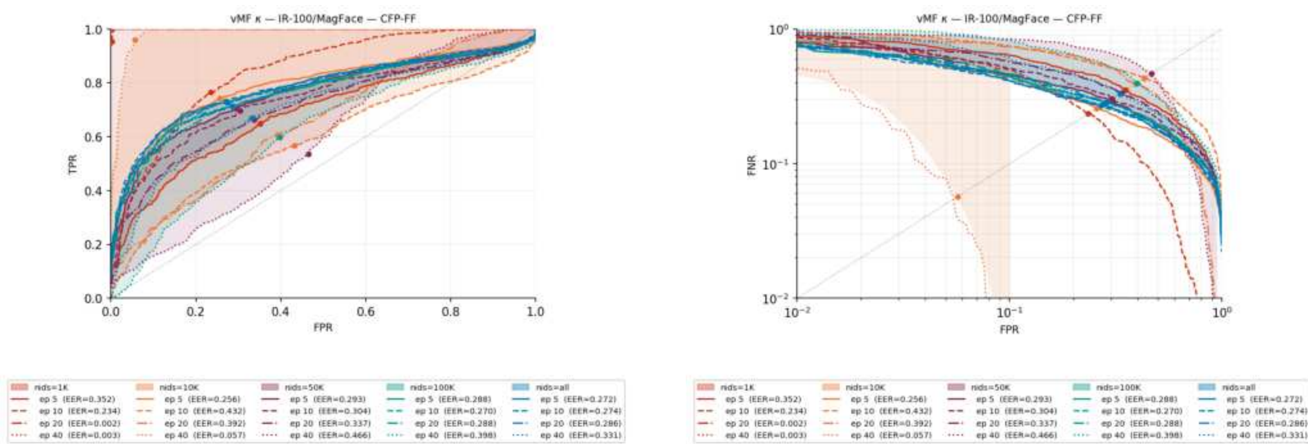


Figure 99. **Threshold-classifier ROC and DET — vMF-κ, CFP-FF.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

CFP-FP

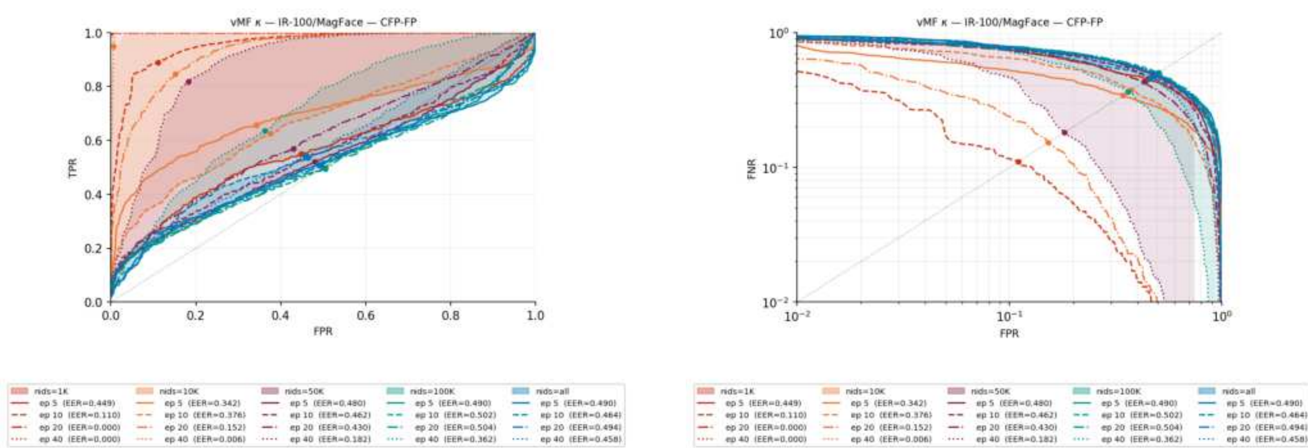


Figure 100. **Threshold-classifier ROC and DET — vMF-κ, CFP-FP.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

CPLFW

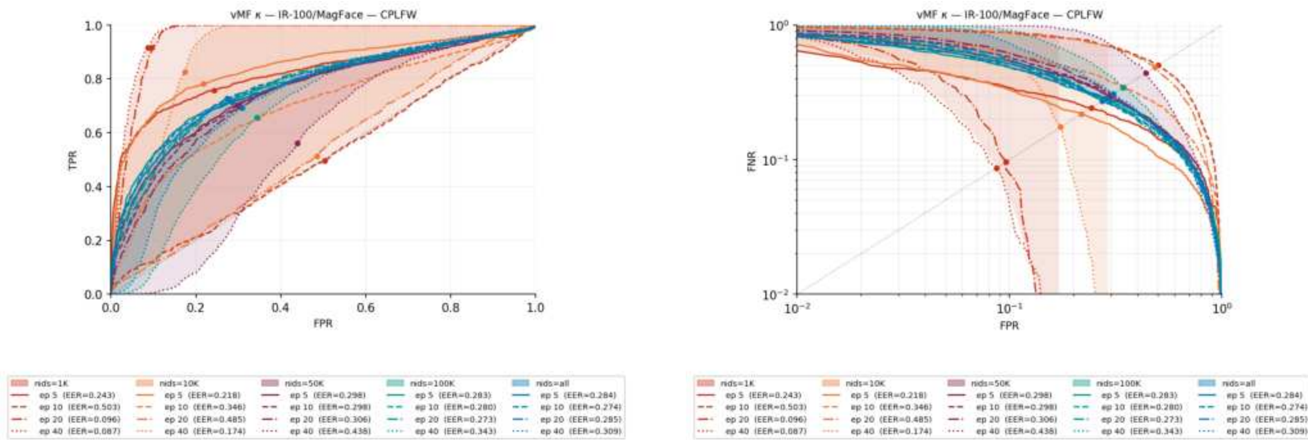


Figure 101. **Threshold-classifier ROC and DET — vMF-κ, CPLFW.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

LFW

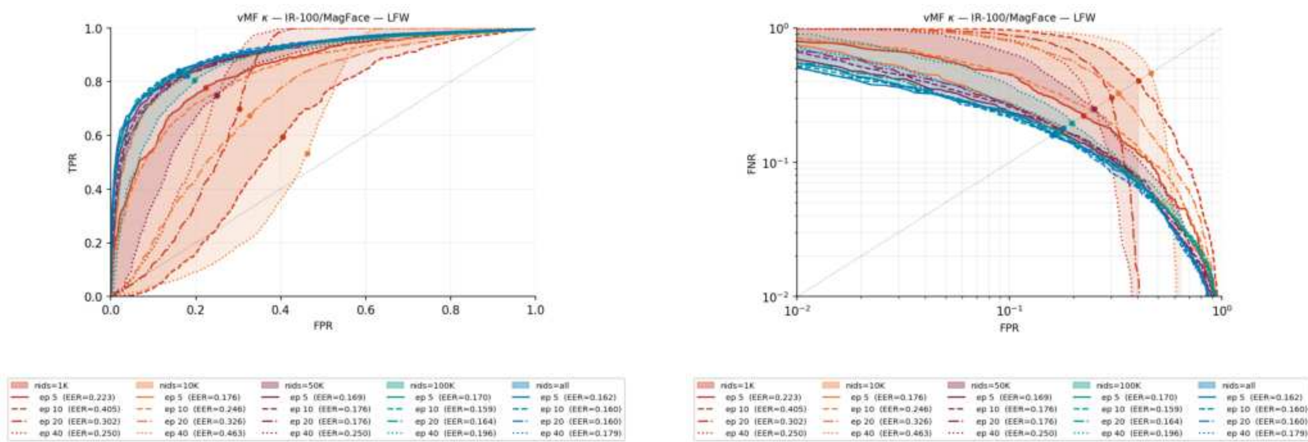


Figure 102. **Threshold-classifier ROC and DET — vMF-κ, LFW.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

RFW

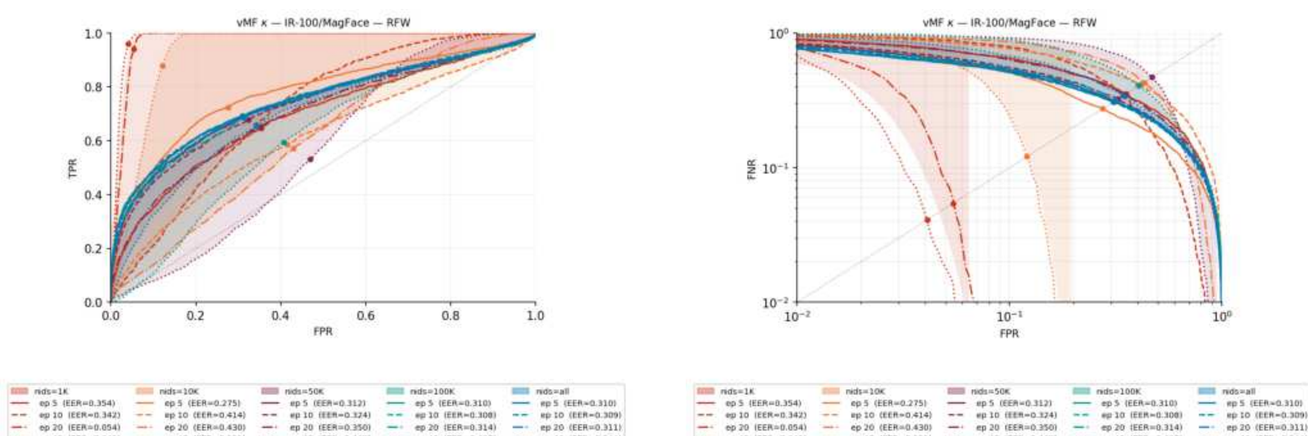


Figure 103. **Threshold-classifier ROC and DET — vMF-κ, RFW.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

XQLFW

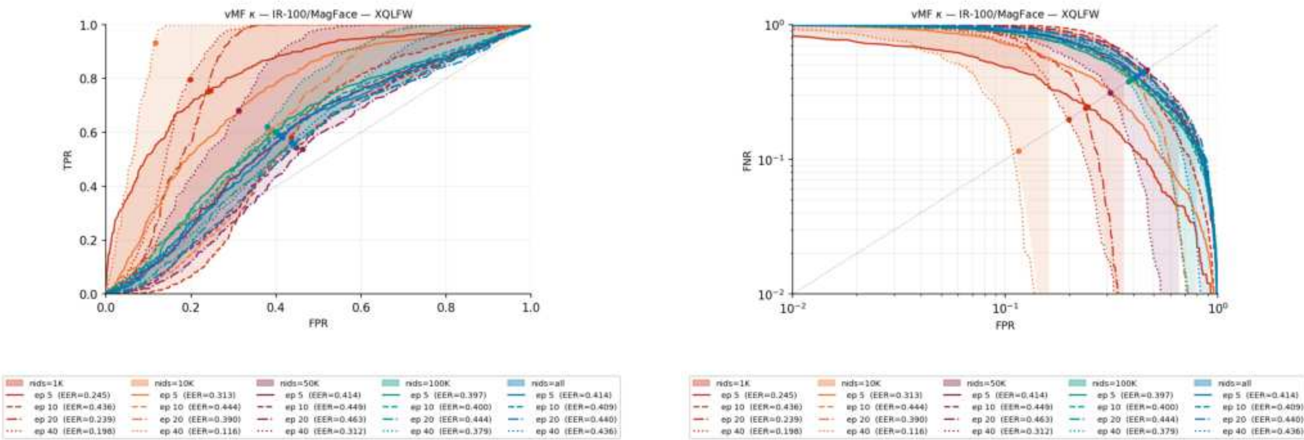


Figure 104. **Threshold-classifier ROC and DET — vMF-κ, XQLFW.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

WebFace4M (held-out)

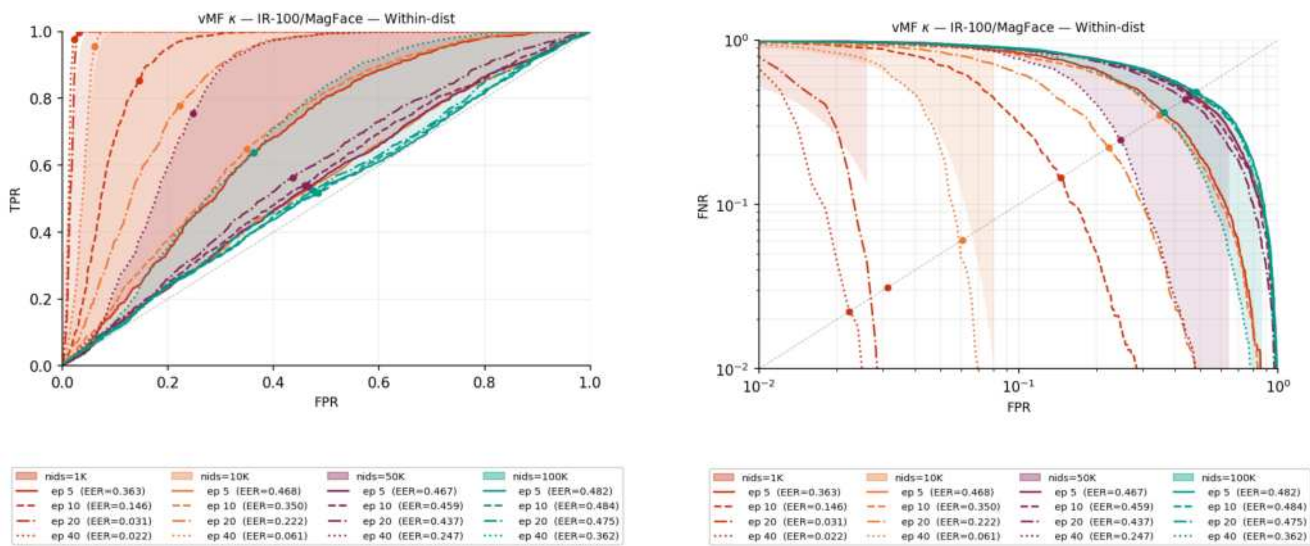


Figure 105. **Threshold-classifier ROC and DET — vMF-κ, WebFace4M (held-out).** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

2.3.3 PenLogit

AgeDB-30

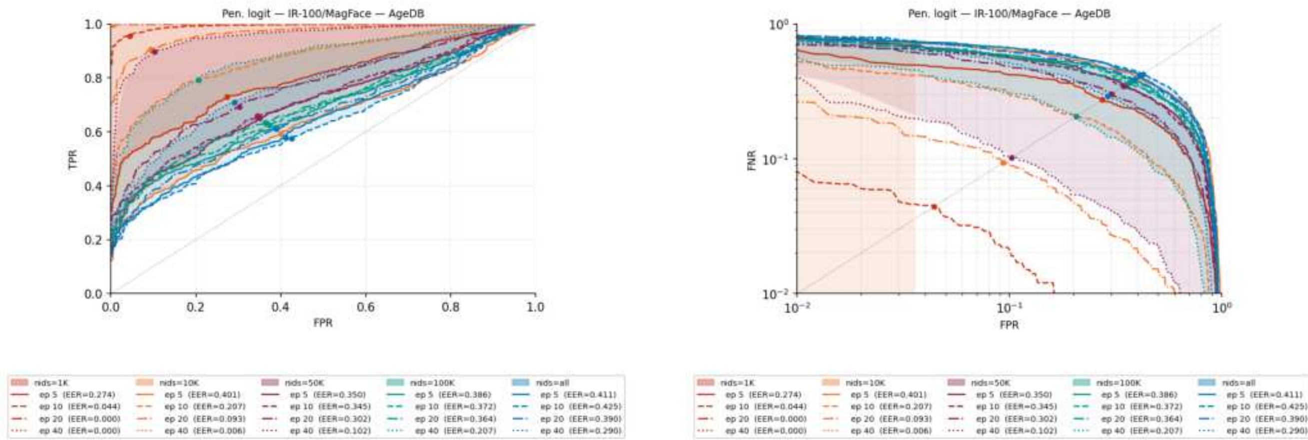


Figure 106. **Threshold-classifier ROC and DET — PenLogit, AgeDB-30.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

CFP-FF

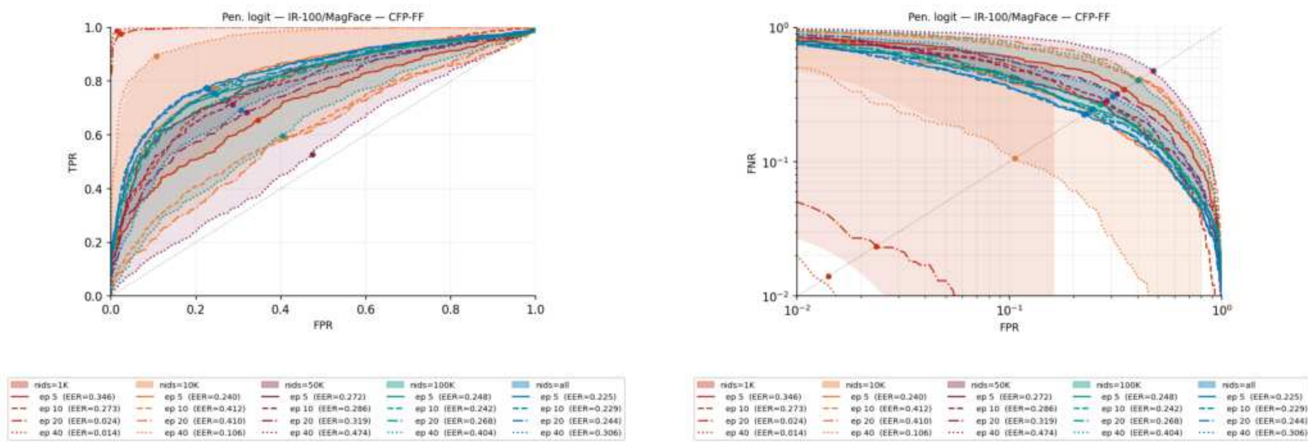


Figure 107. **Threshold-classifier ROC and DET — PenLogit, CFP-FF.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

CFP-FP

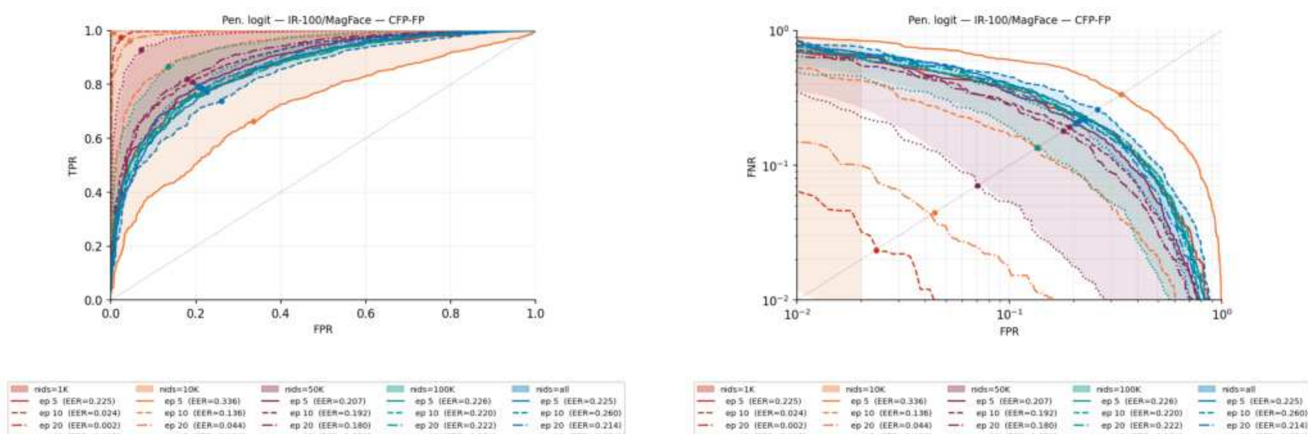


Figure 108. **Threshold-classifier ROC and DET — PenLogit, CFP-FP.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

CPLFW

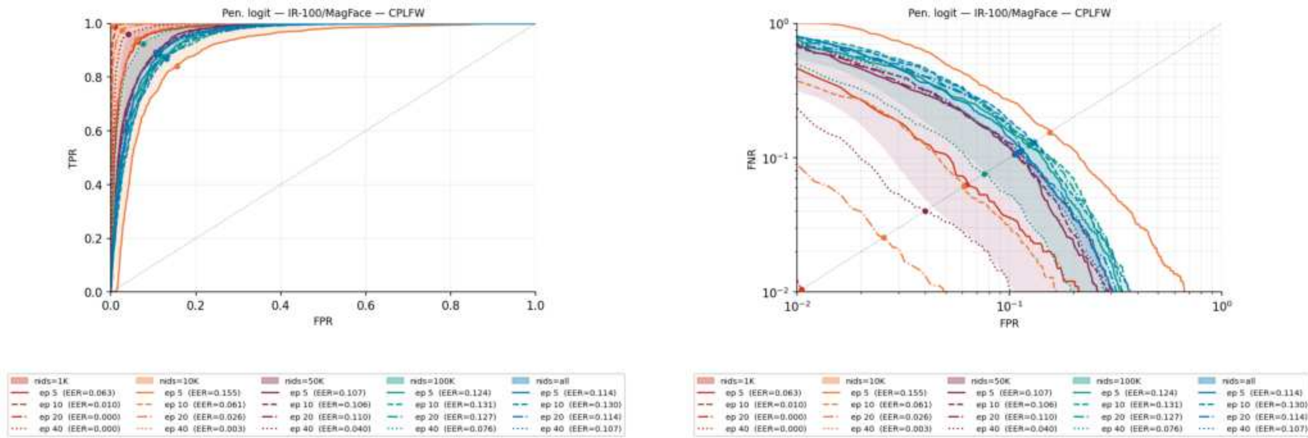


Figure 109. **Threshold-classifier ROC and DET — PenLogit, CPLFW.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

LFW

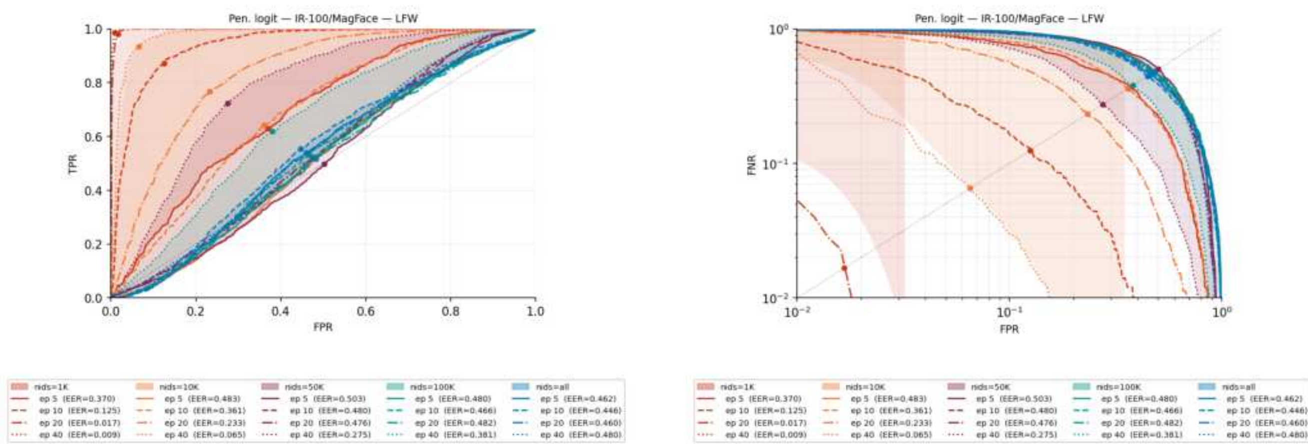


Figure 110. **Threshold-classifier ROC and DET — PenLogit, LFW.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

RFW

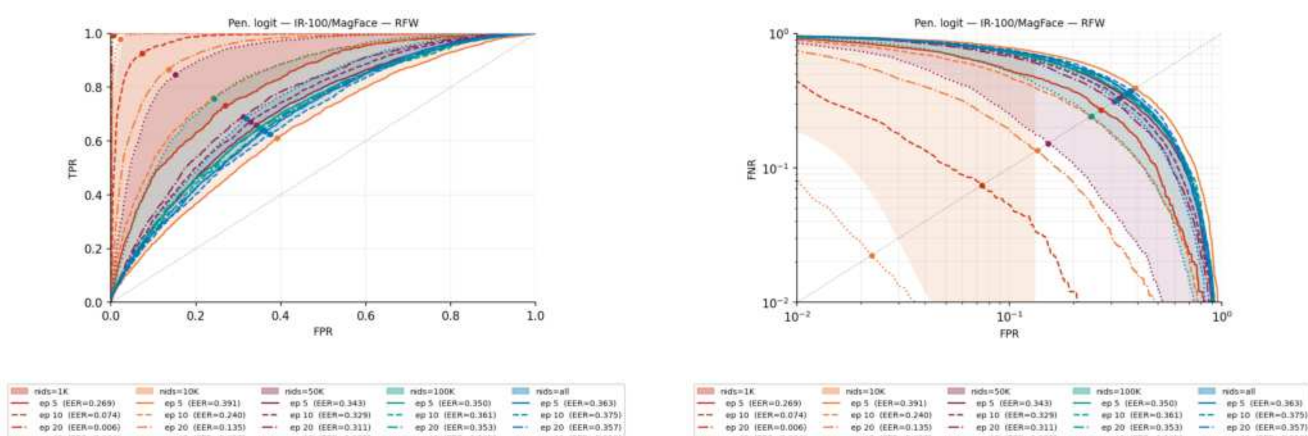


Figure 111. **Threshold-classifier ROC and DET — PenLogit, RFW.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

XQLFW

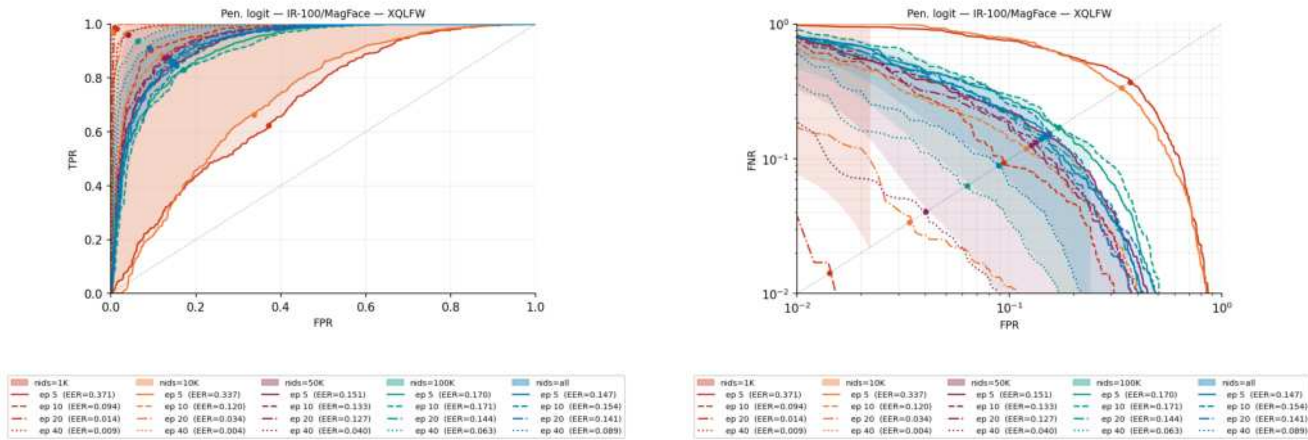


Figure 112. **Threshold-classifier ROC and DET — PenLogit, XQLFW.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

WebFace4M (held-out)

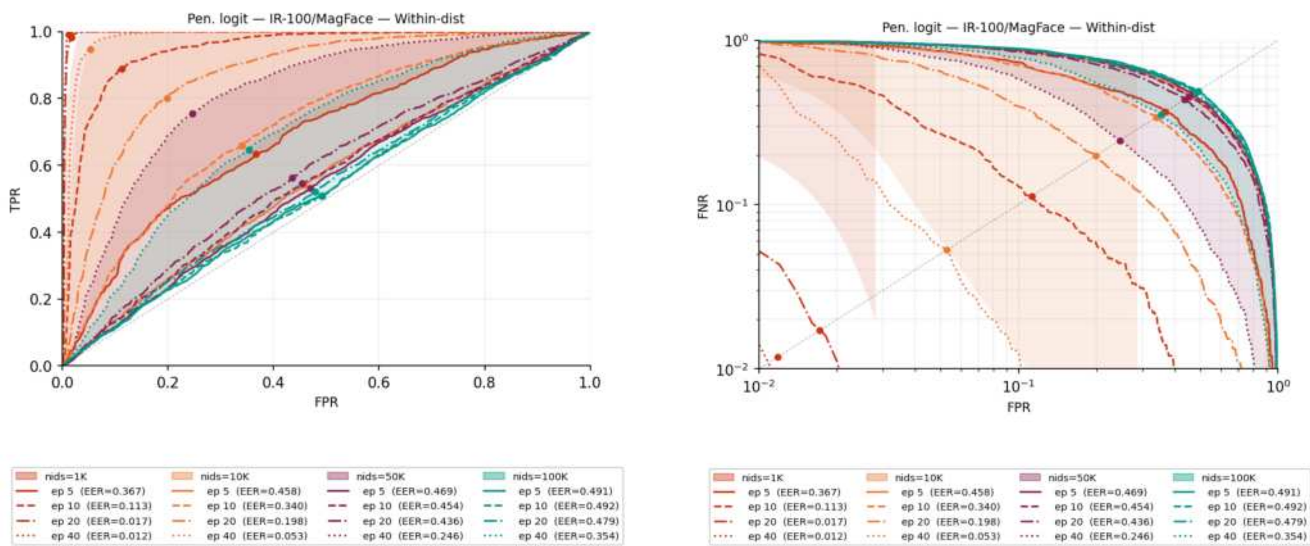


Figure 113. **Threshold-classifier ROC and DET — PenLogit, WebFace4M (held-out).** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

2.3.4 ProtoCE

AgeDB-30

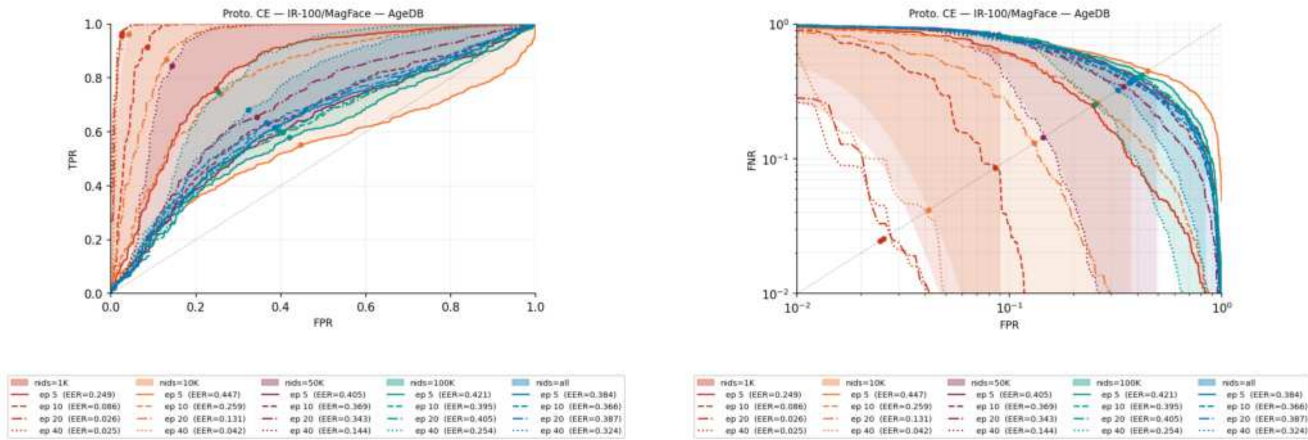


Figure 114. **Threshold-classifier ROC and DET — ProtoCE, AgeDB-30.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

CFP-FF

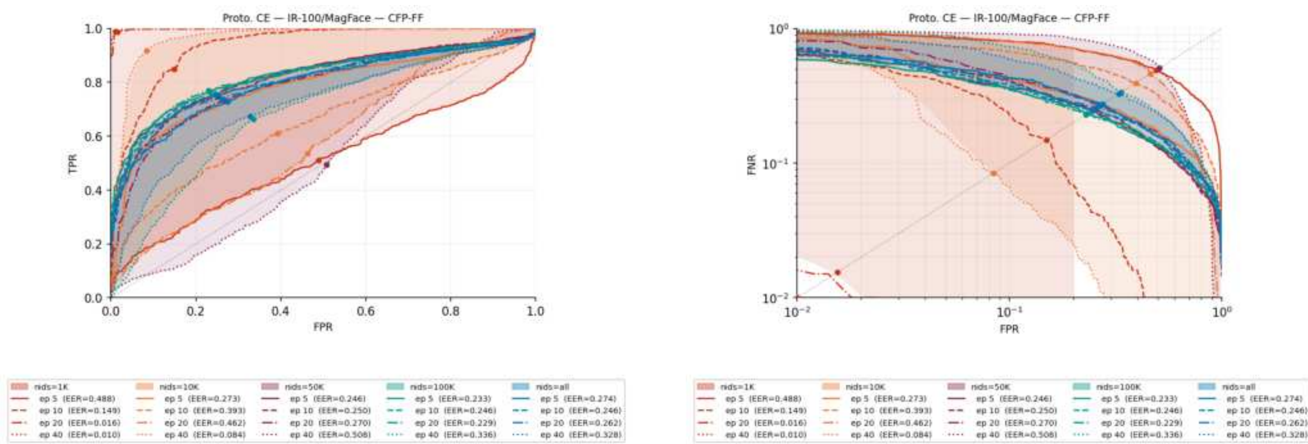


Figure 115. **Threshold-classifier ROC and DET — ProtoCE, CFP-FF.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

CFP-FP

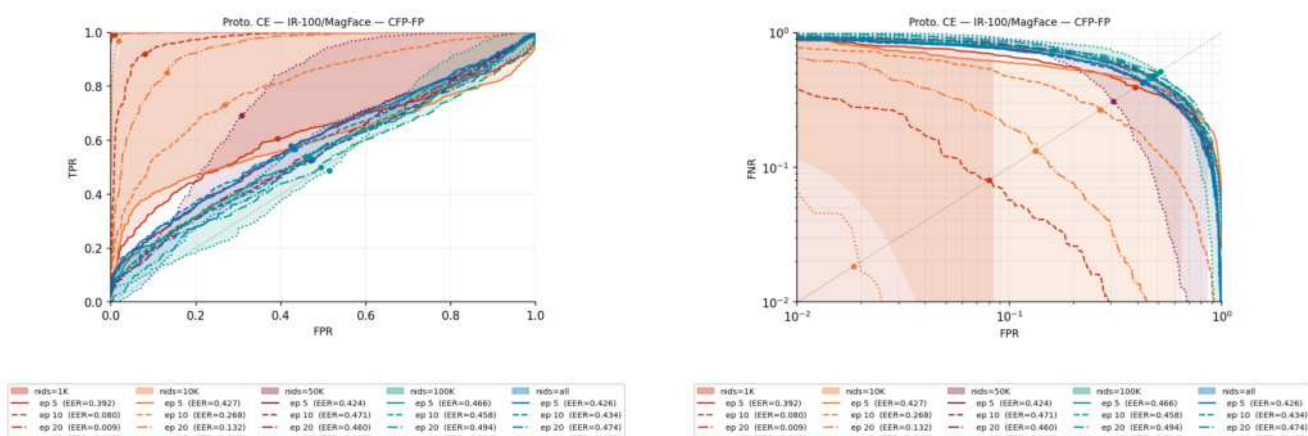


Figure 116. **Threshold-classifier ROC and DET — ProtoCE, CFP-FP.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

CPLFW

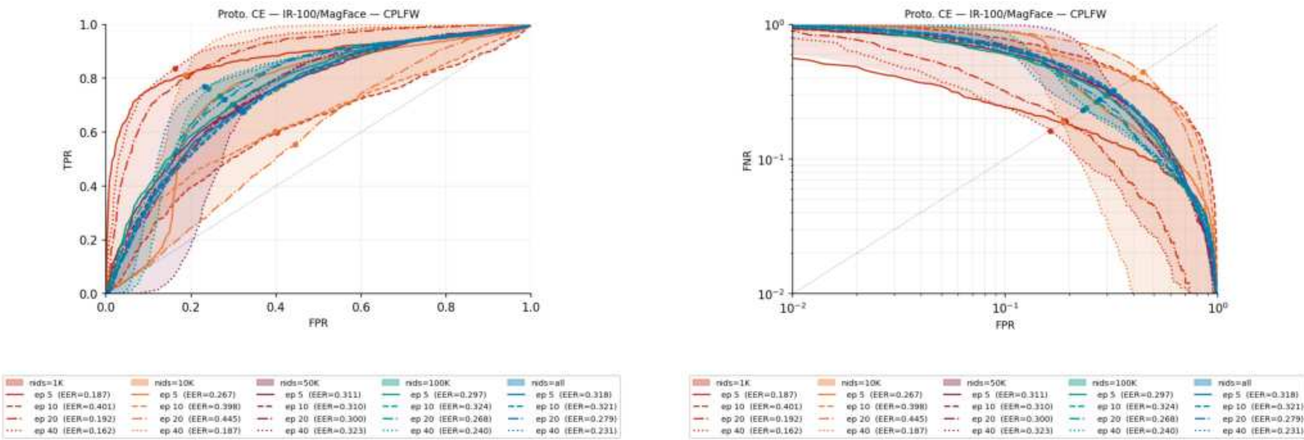


Figure 117. **Threshold-classifier ROC and DET — ProtoCE, CPLFW.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

LFW

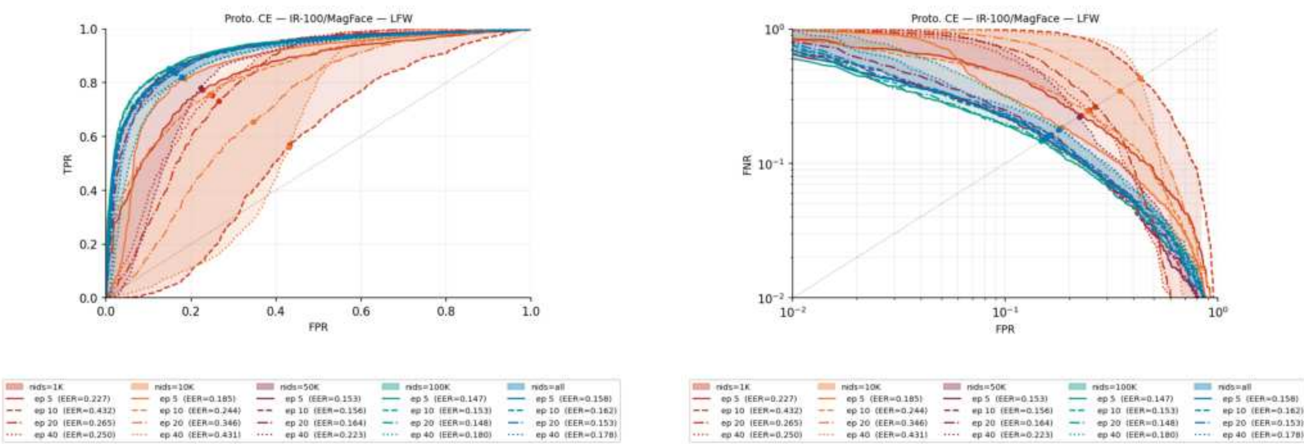


Figure 118. **Threshold-classifier ROC and DET — ProtoCE, LFW.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

RFW

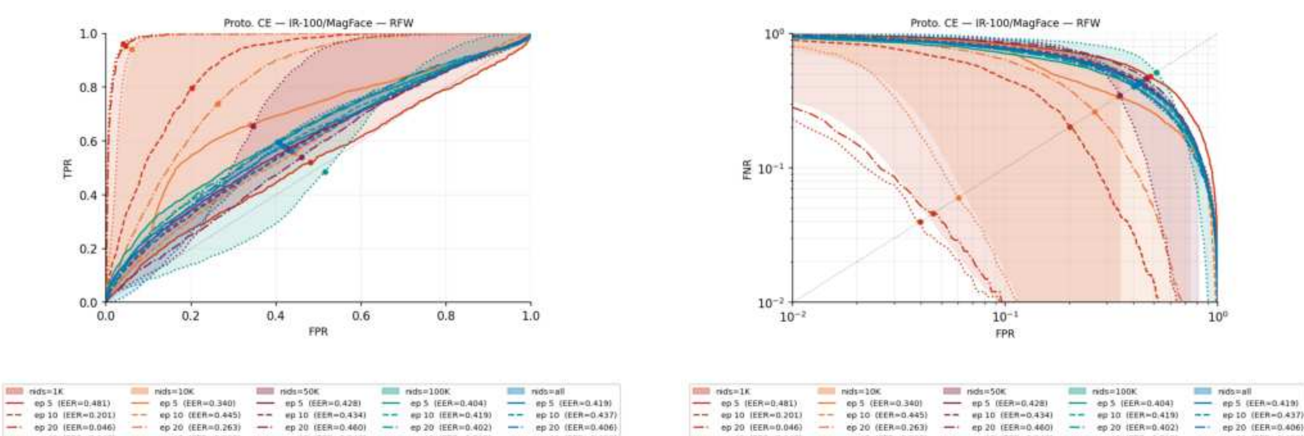


Figure 119. **Threshold-classifier ROC and DET — ProtoCE, RFW.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

XQLFW

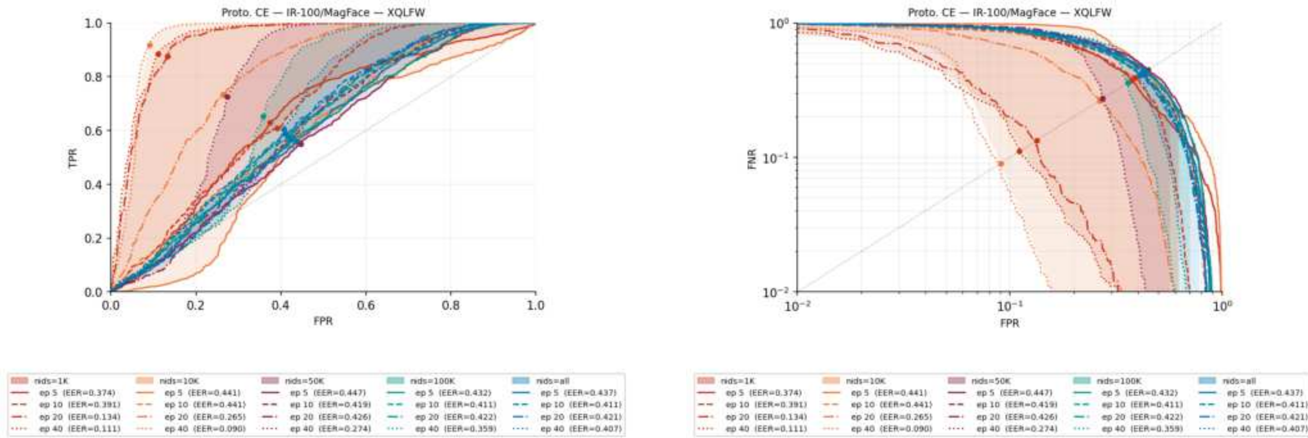


Figure 120. **Threshold-classifier ROC and DET — ProtoCE, XQLFW.** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

WebFace4M (held-out)

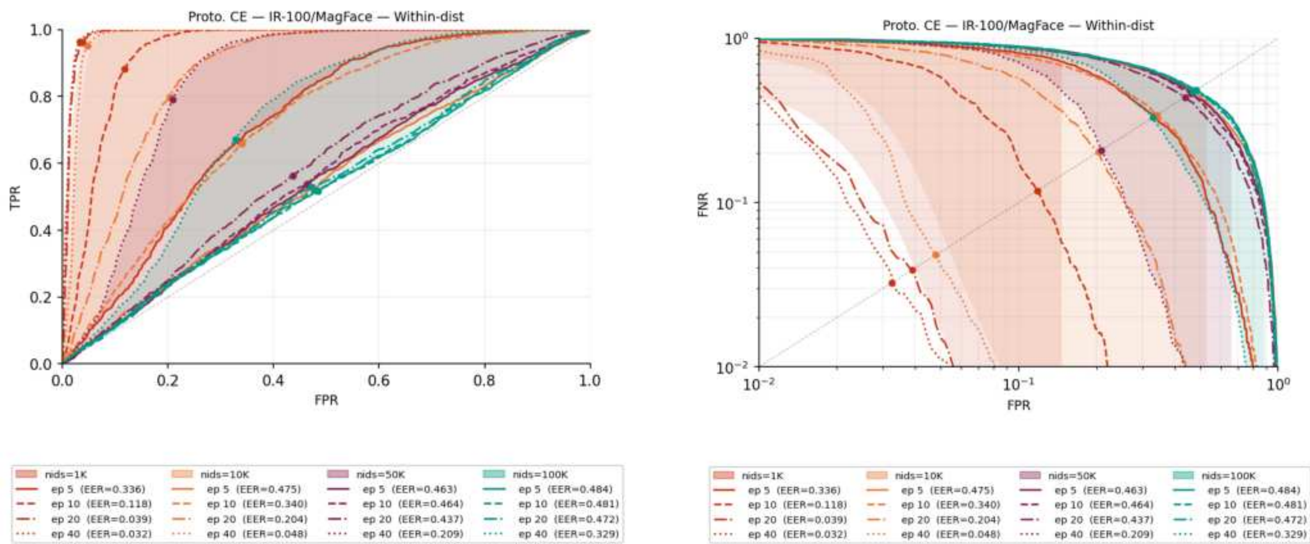


Figure 121. **Threshold-classifier ROC and DET — ProtoCE, WebFace4M (held-out).** Curves are grouped by $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$. ROC (left): true positive rate vs. false positive rate. DET (right): false non-match rate vs. false match rate.

3. Per-epoch verification and membership separability

Each table below reports verification EER (%) and threshold-classifier AUC for all four cluster statistics across every backbone, loss head, training-set size, and evaluation benchmark. **Column conventions.** Training-set size n_{ids}: 1K, 10K, 50K, 100K, All. Sub-rows per backbone$\times$loss block: **E** = EER (%; $\downarrow$), **C** = PairCos AUC, $\boldsymbol{\kappa}$ = vMF-κ AUC, **L** = PenLogit AUC, **S** = ProtoCE AUC. Higher AUC = more recoverable membership information. Evaluation datasets (all cross-domain): LFW, XQLFW, CPLFW, RFW, CFP-FP, CFP-FF, AgeDB, IJB-C.

3.1. Epoch 5

Table 1. **Verification performance and membership separability at epoch 5.** Column conventions as described in Section 3. **Bold**: best EER or AUC per backbone.

Backbone Head		LFW					XQLFW					CPLFW					RFW				
		1K	10K	50K	100K	All	1K	10K	50K	100K	All	1K	10K	50K	100K	All	1K	10K	50K	100K	All
IResNet34 *ArcFace*	E	**10.1**	2.2	0.7	0.5	0.5	22.2	16.5	**12.1**	11.5	12.0	**41.1**	**27.7**	17.7	15.9	15.8	**36.0**	22.1	12.9	11.0	10.3
	C	**.54**	.69	**.76**	.78	.78	.56	**.68**	.77	.74	.75	**.89**	**.83**	.76	.76	**.77**	.68	.56	.51	**.52**	.52
	κ	.84	.88	.91	.91	.91	.84	.70	.62	.65	.65	.83	.77	.79	.79	.78	.67	.71	.75	.75	.75
	L	.72	.58	.50	**.51**	.52	.70	**.87**	.92	.90	.90	**.98**	.97	.95	.95	**.95**	.84	.75	.71	.70	.69
	S	.84	**.90**	.93	.93	**.92**	.63	**.53**	.51	.51	.56	.89	.83	**.82**	**.83**	**.81**	.51	**.58**	**.66**	**.69**	**.67**
IResNet34 *CosFace*	E	10.8	**2.2**	**0.6**	**0.5**	0.5	**22.2**	**15.5**	13.9	11.2	**11.0**	41.3	28.4	**16.9**	15.4	**14.8**	36.5	22.5	**12.2**	10.7	9.6
	C	.53	**.70**	.76	.78	**.79**	**.56**	.64	.71	.73	.73	.89	.81	**.77**	.76	.74	**.68**	.56	**.52**	.51	.51
	κ	**.85**	**.89**	**.91**	**.91**	**.92**	**.85**	**.74**	**.69**	**.68**	**.67**	**.86**	**.83**	**.83**	**.82**	**.83**	**.70**	**.74**	**.75**	**.76**	.76
	L	**.73**	.57	**.52**	.51	**.54**	.69	.84	.89	.89	.90	.98	.97	**.96**	.95	.95	**.84**	.76	**.74**	.71	**.70**
	S	**.84**	.89	**.94**	**.93**	.92	.53	.51	**.61**	**.59**	**.63**	**.92**	**.85**	.79	.79	.76	.51	.56	.62	.66	.64
IResNet34 *MagFace*	E	10.2	2.2	0.6	0.7	**0.5**	22.6	17.4	12.6	**11.0**	12.5	42.1	28.2	17.3	**15.2**	15.1	36.3	**21.8**	12.5	**10.2**	**9.4**
	C	.52	.69	.76	**.78**	.79	.51	.66	**.77**	**.78**	**.75**	.89	.82	.77	**.76**	.76	.67	**.57**	.51	.51	**.53**
	κ	.84	.88	.91	.91	.91	.82	.71	.63	.61	.65	.83	.77	.79	.79	.79	.69	.70	.75	.75	**.76**
	L	.73	**.59**	.51	.51	.53	**.73**	.86	**.92**	**.92**	**.91**	.98	**.97**	.96	**.95**	.95	.83	**.77**	.71	**.71**	.69
	S	.83	.88	.92	.92	.92	**.64**	.51	.56	.54	.60	.88	.79	.73	.75	.75	**.51**	.58	.60	.65	.61
IResNet50 *ArcFace*	E	**10.5**	**1.9**	**0.6**	0.6	**0.5**	23.3	16.6	**12.2**	11.9	10.9	**41.0**	28.3	16.6	14.5	**14.0**	35.9	22.6	12.4	10.1	8.9
	C	.54	**.70**	**.77**	**.78**	**.80**	.55	**.67**	**.79**	.78	**.76**	.88	.81	**.77**	**.76**	**.76**	.67	.56	.50	**.52**	**.54**
	κ	.84	.88	.91	**.91**	**.91**	.84	.70	.60	.61	.62	.83	.78	.78	.78	.79	.68	.71	.74	.75	**.76**
	L	.71	.57	.50	**.51**	**.53**	**.70**	**.86**	**.93**	**.92**	**.91**	.98	**.97**	.96	**.95**	**.95**	.83	.76	.72	.70	.68
	S	.84	.90	**.93**	.93	**.92**	.62	**.54**	.51	.50	.55	.88	.82	**.83**	**.85**	**.82**	.51	**.58**	**.66**	**.69**	**.67**
IResNet50 *CosFace*	E	11.2	2.5	0.7	**0.5**	**0.5**	23.6	**16.4**	12.3	10.7	**9.9**	41.6	**27.7**	**15.8**	**13.9**	14.1	36.7	22.6	**11.4**	**9.8**	**8.7**
	C	.54	.69	.76	.76	.78	**.58**	.62	.73	.75	.76	**.89**	**.82**	.76	.76	.75	**.68**	**.56**	**.52**	.51	.51
	κ	**.85**	**.88**	**.91**	.90	.91	**.86**	**.75**	**.68**	**.66**	**.64**	**.86**	**.82**	**.83**	**.81**	**.82**	**.70**	**.73**	**.75**	**.76**	.76
	L	**.72**	.57	**.51**	.50	.53	.68	.83	.90	.90	.90	**.98**	.97	.96	.95	.95	**.83**	**.76**	**.74**	**.71**	**.70**
	S	**.84**	**.90**	.93	**.93**	.92	.54	.52	**.59**	**.57**	.59	**.92**	**.85**	.82	.84	.80	**.52**	.56	.64	.65	.64
IResNet50 *MagFace*	E	11.0	2.4	0.7	0.6	0.6	**22.0**	17.3	12.4	**10.1**	11.1	41.6	28.8	16.8	13.9	14.4	**35.9**	**22.5**	12.4	10.8	9.5
	C	**.55**	.70	.77	.77	.79	.56	.64	.75	**.78**	.75	.87	.80	.77	.74	.75	.67	.56	.50	.52	.54
	κ	.85	.88	.91	.90	.91	.84	.72	.65	.61	.64	.84	.80	.79	.79	.79	.68	.71	.75	.75	.76
	L	.71	**.57**	.51	.51	.53	.70	.85	.91	.92	.91	.98	.96	**.96**	.95	.95	.82	.76	.72	.70	.68
	S	.83	.88	.92	.92	.91	**.65**	.51	.56	.56	**.62**	.87	.77	.73	.76	.74	.51	.58	.61	.63	.60
IResNet100 *ArcFace*	E	**10.5**	**1.8**	**0.5**	**0.4**	0.5	**21.4**	**14.0**	**11.3**	**10.2**	10.1	40.5	**26.9**	15.9	13.8	13.1	36.3	**21.3**	10.6	9.0	**7.4**
	C	.55	.70	**.77**	**.78**	.79	.58	**.68**	.79	**.81**	.80	.87	.80	.77	.75	**.76**	.66	.57	.51	**.53**	.52
	κ	.85	.88	**.91**	.91	.91	.84	.69	.60	.58	.60	.83	.78	.77	.77	.78	.68	.70	.73	.75	.75
	L	.70	.57	.50	**.52**	.52	.68	**.86**	.92	**.93**	.93	.98	**.96**	.95	.95	**.95**	.82	.76	.72	.69	.70
	S	.84	.90	.93	.93	**.92**	.65	.55	.53	.54	.54	.89	.79	.81	**.83**	.81	**.53**	.58	**.66**	**.68**	**.68**
IResNet100 *CosFace*	E	10.6	2.4	0.6	0.5	0.5	23.4	18.1	11.4	10.8	10.3	41.0	27.3	**15.0**	**12.8**	**12.6**	**35.9**	21.5	**9.9**	**8.3**	7.4
	C	.53	.69	.75	.78	**.79**	**.61**	.62	.78	.79	.74	**.87**	**.80**	.77	.75	.74	**.68**	**.57**	**.53**	.50	.52
	κ	**.85**	.88	.91	**.91**	**.92**	**.87**	**.75**	**.65**	**.63**	**.66**	**.86**	.82	**.81**	**.81**	**.81**	**.70**	.72	**.75**	**.75**	**.76**
	L	**.72**	**.57**	**.52**	.51	**.54**	.65	.83	.92	.92	.89	**.98**	.96	.96	**.95**	.94	**.83**	**.76**	**.74**	**.71**	**.70**
	S	**.85**	**.90**	**.93**	**.93**	.92	.53	**.57**	.57	.58	**.66**	**.91**	**.84**	**.84**	.82	**.81**	.51	.58	.65	.66	.64
IResNet100 *MagFace*	E	10.5	6.8	0.6	0.5	**0.5**	22.4	27.5	11.5	11.3	**9.6**	**40.5**	37.6	15.8	12.9	12.8	36.0	30.3	10.8	8.5	7.5
	C	**.56**	**.75**	.77	.78	.79	.54	.52	**.80**	.78	**.81**	.87	.69	**.78**	**.75**	.76	.64	.56	.50	.51	**.52**
	κ	.85	**.89**	.90	.91	.91	.83	.74	.59	.62	.58	.82	**.85**	.76	.78	.77	.69	**.78**	.74	.75	.75
	L	.69	.51	.51	.51	.53	**.70**	.72	**.93**	.91	**.93**	.98	.91	**.96**	.95	.95	.81	.65	.71	.70	.69
	S	.84	.86	.92	.92	.92	**.65**	.54	**.58**	**.60**	.60	.88	.74	.75	.77	.74	.53	**.68**	.60	.64	.61

Backbone Head		CFP-FP					CFP-FF					AgeDB					IJB-C				
		1K	10K	50K	100K	All	1K	10K	50K	100K	All	1K	10K	50K	100K	All	1K	10K	50K	100K	All
IResNet34 *ArcFace*	E	**28.6**	12.9	5.8	4.5	**4.1**	11.9	2.7	0.8	0.6	0.6	**27.3**	14.2	7.8	6.4	6.4	10.0	**3.9**	**2.5**	**2.0**	2.1
	C	**.79**	**.77**	**.73**	**.74**	.71	.70	.75	.80	.81	.83	.86	.76	.69	.66	**.68**	.51	.51	.53	.51	.54
	κ	.52	.52	.50	.51	.52	.67	.71	.77	.77	.78	.82	.73	.67	.64	**.66**	**.63**	**.60**	**.56**	**.57**	.56
	L	**.89**	**.88**	**.86**	**.87**	.85	.70	.75	.81	.82	.84	.83	.75	.69	.66	**.66**	.60	.59	.62	.60	.63
	S	.66	.65	.56	.52	.60	.55	**.73**	**.83**	**.82**	.80	.83	.69	.60	.55	.58	.68	.64	.66	.65	.71
IResNet34 *CosFace*	E	30.5	13.3	5.5	**3.8**	4.6	12.6	**2.4**	**0.7**	**0.6**	**0.6**	32.0	**14.0**	7.7	**5.8**	**5.4**	10.9	4.2	3.0	2.5	2.4
	C	.70	.73	.69	.70	.69	**.73**	**.77**	**.81**	**.83**	**.83**	**.88**	.76	**.71**	**.68**	.65	**.54**	**.52**	**.58**	**.58**	**.55**
	κ	**.62**	**.55**	**.57**	**.56**	**.56**	**.69**	**.72**	**.77**	**.79**	**.78**	**.82**	.72	**.68**	**.66**	.64	.59	.58	.52	.52	.55
	L	.84	.86	.84	.84	.83	**.73**	**.78**	**.83**	**.85**	**.85**	**.84**	.74	.70	**.67**	.63	**.65**	**.61**	**.67**	**.67**	**.64**
	S	**.71**	**.68**	**.66**	**.61**	**.61**	**.65**	.72	.82	.81	.80	.82	.67	.60	.54	.58	**.71**	**.64**	**.70**	**.72**	**.73**
IResNet34 *MagFace*	E	29.2	**12.7**	**5.4**	4.4	4.1	**11.5**	2.6	0.7	0.7	0.6	31.3	14.3	**7.6**	6.0	5.9	**9.9**	4.1	2.6	2.3	**2.0**
	C	.79	.74	.70	.73	**.72**	.68	.75	.80	.82	.83	.83	**.78**	.70	.65	.65	.50	.50	.54	.53	.53
	κ	.52	.51	.54	.51	.51	.64	.71	.76	.78	.78	.79	**.75**	.68	.63	.64	.62	.59	.56	.56	**.57**
	L	.89	.87	.84	.86	**.86**	.68	.76	.81	.83	.84	.80	**.77**	**.70**	.65	.64	.61	.60	.63	.61	.63
	S	.63	.67	.64	.57	.60	.53	.72	.82	.82	**.80**	**.84**	**.69**	**.61**	**.58**	**.63**	.66	.63	.66	.63	.67
IResNet50 *ArcFace*	E	29.0	**11.8**	5.7	3.8	3.8	12.4	**2.2**	0.8	0.6	**0.6**	29.8	14.7	7.0	5.9	5.6	**9.9**	**4.0**	**2.5**	**2.0**	2.0
	C	.74	.73	**.76**	**.77**	.75	**.72**	.75	.80	**.81**	**.83**	.85	.76	.71	.65	.66	.50	.51	.50	.50	.51
	κ	**.56**	.52	.53	**.54**	.52	**.69**	.71	.76	**.77**	**.78**	.81	.73	.68	.63	.64	**.62**	**.59**	**.58**	**.58**	**.58**
	L	.86	.86	**.88**	**.88**	.87	**.72**	.75	.81	**.82**	.84	.82	.74	**.70**	.65	.64	.61	.60	.59	.59	.61
	S	**.65**	.65	.55	.51	.52	.54	**.73**	.82	.82	**.80**	.83	**.68**	.59	.56	.59	.68	.64	.68	.67	.70
IResNet50 *CosFace*	E	29.7	13.5	**4.4**	4.1	**3.8**	12.6	2.6	0.7	**0.4**	0.6	31.1	**14.4**	**7.0**	**5.5**	5.4	10.7	4.3	2.6	2.1	1.9
	C	.78	.72	.72	.75	.72	.70	**.76**	**.81**	.81	.83	**.88**	.77	**.71**	.66	**.68**	**.52**	**.54**	**.57**	**.55**	**.55**
	κ	.55	**.55**	**.55**	.51	.53	.66	**.72**	**.76**	.76	.78	**.82**	.73	**.68**	.64	**.66**	.61	.56	.53	.54	.55
	L	.89	.85	.85	.88	.85	.70	**.77**	**.82**	.82	**.84**	**.84**	.75	.70	.65	**.66**	**.63**	**.63**	**.66**	**.64**	**.64**
	S	.63	**.69**	.60	.52	.52	**.63**	.72	.81	.82	.78	.82	.65	.59	.57	.59	**.70**	**.65**	**.72**	**.70**	**.74**
IResNet50 *MagFace*	E	**28.5**	12.7	4.6	**3.2**	3.9	**11.6**	2.7	**0.7**	0.6	**0.6**	**28.6**	**14.4**	7.5	5.8	**5.4**	10.3	4.1	2.7	2.0	**1.9**
	C	**.79**	**.74**	.72	.75	**.76**	.68	.74	.79	.80	.83	.85	**.78**	.70	**.69**	.66	.52	.52	.55	.53	.53
	κ	.52	.51	.51	.51	**.53**	.65	.71	.76	.77	.78	.81	**.74**	.68	**.66**	.64	.60	.58	.55	.56	.57
	L	**.89**	**.87**	.86	.87	**.88**	.68	.75	.80	.81	.84	.82	**.76**	.69	**.68**	.65	.62	.61	.64	.61	.62
	S	.62	.66	**.63**	**.55**	**.59**	.51	.72	**.82**	**.83**	.80	**.84**	.68	**.63**	**.59**	**.63**	.66	.62	.67	.64	.67
IResNet100 *ArcFace*	E	**28.3**	**11.5**	**3.9**	**3.0**	3.2	**11.3**	**2.1**	0.8	0.5	**0.4**	**29.2**	**13.6**	6.6	5.3	**4.5**	10.2	**3.8**	2.3	1.9	1.9
	C	.73	**.73**	.74	.73	**.76**	.69	.73	**.81**	.81	.82	**.85**	**.79**	**.74**	.68	.64	.53	.50	.53	.51	.50
	κ	.56	.51	.51	.50	.53	.66	.70	**.77**	**.78**	.77	**.80**	**.76**	**.71**	.65	.63	.59	**.59**	.56	.57	**.59**
	L	.86	**.85**	.86	.86	**.87**	.69	.74	.82	.82	.83	**.82**	**.77**	**.72**	.67	.63	.63	.59	.62	.60	.60
	S	.62	.68	.57	.50	.51	.50	.73	.80	.81	**.80**	.81	**.68**	.60	.58	.59	.66	.64	.70	.69	.72
IResNet100 *CosFace*	E	29.3	13.7	4.2	**3.0**	2.7	12.5	2.8	0.8	**0.4**	0.5	30.7	14.2	6.7	**5.1**	4.9	10.1	4.1	2.4	2.1	1.9
	C	.74	.70	.70	.72	.69	.71	.76	.80	**.82**	.83	.84	.77	.72	**.69**	**.68**	**.53**	.53	**.55**	**.54**	**.55**
	κ	**.59**	.56	**.55**	**.53**	**.55**	.68	.72	.76	.77	.78	.79	.73	.69	**.66**	**.66**	**.60**	.57	.55	.55	.55
	L	**.86**	.84	.84	.85	.83	**.71**	.77	**.82**	**.83**	**.84**	.81	.75	.70	.68	**.66**	**.64**	.62	**.64**	**.62**	**.64**
	S	**.66**	**.71**	**.64**	.51	.58	**.60**	.72	.79	.82	.78	.81	.64	.58	.54	.60	**.68**	**.65**	**.74**	**.74**	**.76**
IResNet100 *MagFace*	E	28.5	19.5	4.2	3.1	**2.6**	12.2	7.2	**0.7**	0.5	0.5	31.1	23.3	**6.1**	5.2	4.7	**10.1**	8.4	**2.2**	**1.9**	**1.8**
	C	**.74**	.53	**.76**	**.73**	.72	**.72**	**.82**	.80	.81	**.83**	.83	.64	.72	.68	.65	.52	**.65**	.50	.51	.50
	κ	.56	**.69**	.53	.50	.51	**.68**	**.79**	.76	.77	**.78**	.79	.62	.69	.66	.64	.59	.54	**.59**	**.57**	.58
	L	.86	.72	**.88**	**.86**	.85	.71	**.83**	.81	.82	.84	.80	.65	.71	**.68**	.64	.62	**.73**	.59	.60	.60
	S	.62	.61	.60	**.56**	**.60**	.51	**.78**	**.82**	**.83**	.79	**.82**	.54	**.62**	**.59**	**.63**	.64	.54	.66	.65	.68

3.2. Epoch 10

Table 2. **Verification performance and membership separability at epoch 10.** Column conventions as described in Section 3. **Bold**: best EER or AUC per backbone.

Backbone / Head		LFW					XQLFW					CPLFW					RFW				
		1K	10K	50K	100K	All	1K	10K	50K	100K	All	1K	10K	50K	100K	All	1K	10K	50K	100K	All
IResNet34 *ArcFace*	E	**8.9**	**1.5**	0.6	0.4	**0.4**	22.5	**14.5**	**10.6**	**9.2**	**10.4**	40.0	24.1	15.4	14.7	13.9	**34.6**	19.8	11.2	10.5	**8.8**
	C	.86	**.59**	**.77**	**.79**	.79	.87	.86	.77	.76	.78	.99	.88	.75	.73	**.75**	.97	.68	.50	.51	.53
	κ	.59	.84	**.91**	**.91**	.91	.58	.54	.62	.63	.62	.51	.71	.79	.80	.79	**.76**	.62	.74	.75	.75
	L	.96	.69	.50	**.53**	.54	.96	.96	.91	.91	.92	**1.**	.98	.95	.94	**.95**	.99	.84	.71	**.70**	.68
	S	.57	.87	.93	.93	**.92**	.70	.54	.52	.50	.53	.66	.72	.79	.81	.80	.89	.55	**.65**	**.68**	**.68**
IResNet34 *CosFace*	E	10.1	1.6	**0.5**	0.4	0.4	24.1	17.2	10.8	10.4	11.4	**39.0**	24.1	**14.7**	**14.2**	**13.7**	34.8	19.8	**10.5**	10.5	8.9
	C	**.86**	.59	.76	.78	.80	.87	.82	**.78**	.73	**.80**	.99	**.88**	**.78**	**.74**	.75	**.97**	**.68**	.52	.52	.53
	κ	**.63**	**.85**	.91	.91	**.92**	**.61**	**.63**	**.65**	**.66**	**.62**	**.62**	**.78**	**.82**	**.82**	**.81**	.71	**.66**	**.75**	**.76**	**.76**
	L	**.96**	**.70**	**.52**	.52	**.55**	.96	.94	**.92**	.89	.92	1.	**.99**	**.96**	**.95**	.95	**.99**	**.85**	.73	.70	**.69**
	S	**.59**	**.89**	**.93**	**.93**	.91	**.77**	**.61**	.56	**.60**	.57	**.79**	**.81**	**.85**	**.82**	**.82**	.87	.56	.63	.63	.64
IResNet34 *MagFace*	E	9.3	1.6	0.6	**0.4**	**0.4**	**21.7**	15.1	12.2	9.6	10.9	39.5	**23.3**	15.3	14.4	13.8	35.0	**19.6**	10.5	**10.2**	9.2
	C	.86	.58	.76	.78	**.80**	**.90**	**.86**	.76	**.78**	.79	**.99**	.88	.76	.73	.74	.96	.67	**.53**	**.52**	**.53**
	κ	.59	.83	.90	.91	.91	.54	.53	.63	.61	.60	.52	.69	.79	.79	.79	.74	.62	.73	.75	.75
	L	.96	.69	.52	.53	.54	**.97**	**.96**	.92	**.92**	**.92**	1.	.98	.96	.94	.95	.98	.83	**.74**	.69	.68
	S	.54	.84	.92	.92	.92	.71	.58	**.61**	.57	**.59**	.62	.66	.72	.73	.74	**.90**	**.58**	.61	.63	.61
IResNet50 *ArcFace*	E	**9.0**	**1.4**	**0.5**	**0.4**	0.4	21.7	**14.0**	**9.8**	**9.0**	**8.6**	40.0	**22.9**	**13.9**	13.3	12.9	**34.4**	19.5	**10.1**	8.7	**7.7**
	C	.83	**.60**	.76	.79	**.80**	.89	.85	**.81**	.77	**.81**	.99	.86	**.78**	.75	.75	.95	.69	.53	.51	.53
	κ	.60	.84	.90	.91	.92	.55	.53	.58	.61	.57	.50	.70	.77	.78	.78	.73	.61	.73	.74	.75
	L	.95	.67	.52	**.53**	**.54**	.96	.95	**.93**	**.91**	**.93**	1.	.98	**.96**	.95	**.95**	.98	.84	.73	.70	.68
	S	.56	.84	**.93**	**.93**	**.92**	.69	.58	.51	.52	.51	.66	.69	.81	.82	**.83**	.88	.58	**.65**	**.69**	**.67**
IResNet50 *CosFace*	E	10.0	1.7	0.6	0.4	**0.3**	22.8	16.6	9.9	9.1	10.5	40.6	23.8	14.4	**13.1**	**12.2**	35.5	**19.0**	10.3	**8.3**	7.7
	C	**.85**	.58	.75	.78	.79	.85	.81	.75	**.77**	.79	**.99**	.88	.77	**.76**	**.76**	**.97**	.71	**.53**	.51	.51
	κ	**.63**	**.84**	.91	.91	**.92**	**.62**	**.63**	**.66**	**.64**	**.63**	**.61**	**.77**	**.81**	**.81**	**.81**	.71	**.63**	**.74**	**.76**	**.75**
	L	**.96**	**.71**	**.52**	.52	.54	.95	.94	.90	.91	.92	**1.**	.98	.95	**.95**	.95	**.99**	**.86**	**.74**	**.71**	**.70**
	S	**.59**	**.88**	.93	.93	.91	**.78**	.64	.57	.57	.57	**.80**	**.76**	**.82**	**.86**	.81	.88	.57	.63	.67	.65
IResNet50 *MagFace*	E	9.4	1.6	**0.5**	0.5	0.4	**21.0**	15.0	11.0	9.4	10.5	**39.6**	23.8	14.6	13.5	12.4	34.9	19.6	10.4	8.9	8.7
	C	.85	.57	**.77**	**.79**	.79	**.89**	**.88**	.78	.76	.76	.99	**.90**	.77	.73	.74	.96	**.71**	.51	**.52**	**.53**
	κ	.59	.82	**.91**	**.91**	.91	.55	.51	.61	.63	.62	.53	.67	.78	.80	.79	**.74**	.59	.74	.75	.75
	L	.95	.70	.51	.52	.54	**.97**	**.96**	.92	.91	.91	1.	**.99**	.96	.94	.95	.98	.85	.73	.70	.69
	S	.54	.82	.92	.92	.92	.68	**.64**	**.59**	**.58**	**.62**	.60	.65	.73	.75	.76	**.90**	**.61**	.59	.63	.61
IResNet100 *ArcFace*	E	9.4	**1.4**	0.5	**0.3**	**0.3**	**21.6**	**14.6**	**9.3**	**8.4**	**9.0**	**38.8**	**21.6**	**13.2**	12.0	12.9	35.0	**18.4**	8.9	**7.8**	**6.3**
	C	**.82**	.58	**.75**	.78	.80	.85	**.87**	.83	**.80**	.80	.99	.88	.77	**.74**	**.74**	.94	.69	.53	.50	.53
	κ	.60	.83	.90	.91	.91	.57	.52	.56	.59	**.59**	.53	.68	.76	.78	.78	.71	.61	.73	.74	**.75**
	L	**.94**	**.69**	.52	.51	.54	.95	**.96**	.94	**.93**	.92	1.	.98	.96	.95	**.95**	.98	.84	.73	.71	.68
	S	.58	.86	**.93**	**.93**	**.92**	.70	.60	.54	.54	.56	.66	.69	**.82**	**.82**	**.81**	**.88**	.58	**.64**	**.70**	**.68**
IResNet100 *CosFace*	E	9.7	1.6	**0.5**	**0.3**	0.3	22.9	14.8	11.3	9.6	9.7	40.3	22.9	13.3	**11.8**	**11.7**	35.0	18.5	**8.2**	7.9	6.5
	C	.81	**.59**	.75	.78	.80	.81	.84	.82	.78	**.82**	.99	**.88**	.77	.74	.74	.95	**.70**	**.53**	.50	.52
	κ	**.66**	**.84**	**.90**	.91	**.92**	**.66**	**.60**	**.61**	**.62**	.58	**.64**	**.74**	**.81**	**.81**	**.79**	.66	**.62**	**.74**	**.75**	.75
	L	.94	.69	.53	.51	**.56**	.93	.94	.93	.92	**.93**	1.	.98	.96	**.95**	.94	**.98**	**.85**	**.75**	**.71**	**.69**
	S	**.61**	**.86**	.93	.92	.90	**.78**	**.65**	**.64**	**.64**	.60	**.79**	**.74**	.81	.78	.80	.86	**.61**	.63	.65	.64
IResNet100 *MagFace*	E	**8.8**	1.5	0.5	0.4	0.3	22.2	15.1	11.0	9.2	9.1	39.7	23.5	13.5	12.8	12.0	**34.9**	19.3	9.0	8.5	7.2
	C	.82	.58	.75	**.79**	**.80**	**.87**	.85	**.83**	.76	.79	**.99**	.88	**.78**	.73	.72	**.95**	.68	.52	**.52**	**.53**
	κ	.61	.83	.90	**.91**	.92	.56	.54	.54	.61	.59	.50	.70	.75	.77	.78	**.72**	.61	.73	.74	.75
	L	.94	.69	**.53**	**.53**	.55	**.96**	.95	**.94**	.90	.92	**1.**	**.98**	**.96**	.94	.94	.98	.84	.73	.69	.68
	S	.56	.84	.92	.92	.91	.68	.62	.61	.63	**.63**	.61	.64	.75	.73	.74	.88	.59	.59	.62	.59

Backbone / Head		CFP-FP					CFP-FF					AgeDB					IJB-C				
		1K	10K	50K	100K	All	1K	10K	50K	100K	All	1K	10K	50K	100K	All	1K	10K	50K	100K	All
IResNet34 *ArcFace*	E	**26.9**	10.5	**3.5**	4.0	4.0	**9.9**	2.2	0.7	0.7	0.5	**27.1**	12.2	6.0	6.2	5.2	9.6	3.9	2.2	2.1	2.0
	C	1.	.87	.69	.72	.73	**.91**	.64	.80	**.82**	.83	1.	.89	.74	**.70**	.63	.86	.61	**.54**	.53	.50
	κ	.97	.65	**.54**	.50	.50	**.91**	.62	**.77**	**.78**	.79	**.97**	.85	.71	**.68**	.62	.91	.70	.55	.55	**.59**
	L	**1.**	.94	.83	.85	.85	**.87**	.65	.81	**.83**	.85	.99	.86	.72	**.70**	.62	.77	.51	**.63**	.62	.59
	S	**.98**	**.79**	.59	.51	.55	.96	.65	.81	.81	.80	.96	**.80**	.62	.56	.59	.94	.77	.70	.67	.70
IResNet34 *CosFace*	E	28.9	11.0	4.5	4.2	4.0	11.0	1.8	0.8	0.6	0.5	27.7	11.9	**5.9**	**5.6**	5.1	9.8	4.4	**2.2**	2.2	2.1
	C	1.	.87	.75	.70	**.73**	.88	**.66**	**.80**	.81	**.84**	**1.**	**.89**	**.74**	.70	.63	.84	.58	.51	**.55**	**.51**
	κ	.95	.62	.51	**.54**	**.52**	.88	**.63**	.76	.77	**.79**	.96	.84	.71	.67	.62	.90	.67	**.58**	.54	.58
	L	1.	.94	.88	.84	**.86**	.84	**.67**	**.82**	.82	**.85**	**1.**	.86	.73	.69	.61	.76	**.52**	.60	**.63**	**.60**
	S	.98	.76	.54	.53	.51	**.97**	**.68**	.81	.82	.78	.95	.78	.56	.54	.59	**.95**	**.77**	**.71**	**.70**	**.73**
IResNet34 *MagFace*	E	27.6	**9.7**	3.9	**3.7**	**3.5**	10.3	**1.7**	**0.7**	**0.5**	**0.4**	28.0	**11.6**	6.2	5.8	**5.0**	**9.1**	**3.4**	2.3	**1.9**	**2.0**
	C	**1.**	**.89**	**.76**	**.73**	.72	.91	.64	.79	.80	.83	1.	.89	.74	.70	**.66**	**.86**	**.62**	.53	.51	.51
	κ	**.97**	**.69**	.52	.50	.51	.91	.61	.75	.76	.78	.97	**.85**	**.72**	.67	**.65**	**.91**	**.70**	.56	**.57**	.58
	L	1.	**.94**	**.88**	**.86**	.85	.87	.65	.80	.81	.84	.99	**.86**	**.73**	.69	**.65**	**.77**	.52	.62	.60	.60
	S	.98	.78	**.60**	**.54**	**.59**	.96	.65	**.82**	**.82**	**.81**	**.96**	.80	**.64**	**.60**	**.62**	.93	.74	.67	.63	.66
IResNet50 *ArcFace*	E	**25.9**	9.3	**3.5**	3.3	**3.0**	10.6	1.9	**0.6**	**0.4**	**0.3**	**27.6**	11.2	5.7	5.2	4.8	9.1	**3.3**	**1.9**	**1.8**	**1.9**
	C	.99	.88	**.77**	**.76**	.73	**.89**	.62	.78	.81	.83	1.	.88	**.76**	**.69**	.65	.85	.63	.51	.50	.51
	κ	.95	.70	**.54**	**.53**	.51	**.89**	.59	.75	.77	.78	**.97**	.85	**.73**	**.67**	.65	.90	.71	**.60**	**.58**	**.58**
	L	1.	.94	**.88**	**.87**	.86	**.86**	.63	.79	.81	.84	.99	.86	**.74**	**.68**	.64	.76	.54	.58	.59	.60
	S	**.98**	.77	.56	.53	.52	.95	.63	**.81**	.81	.79	**.96**	.81	.61	.57	.61	.94	**.77**	.70	.67	.72
IResNet50 *CosFace*	E	29.8	11.4	4.2	3.4	3.2	11.0	1.9	0.6	**0.4**	**0.3**	27.7	12.8	**5.7**	**4.8**	**4.8**	9.8	4.2	2.2	2.1	2.1
	C	.99	.88	.74	.75	.74	.86	**.65**	.79	**.81**	**.84**	**1.**	.89	.75	.69	.64	.84	.59	**.53**	**.52**	**.54**
	κ	.94	.65	.52	.50	.51	.87	**.62**	.75	.77	**.79**	.96	.84	.71	.66	.63	.89	.68	.57	.57	.56
	L	1.	.94	.86	.87	.86	.82	**.66**	.80	**.83**	**.86**	**1.**	.86	.73	.68	.63	.75	.51	**.62**	**.61**	**.63**
	S	.98	.79	.55	.51	.54	**.97**	**.66**	.80	.81	.77	.95	.78	.59	.54	.59	**.95**	.76	**.72**	**.72**	**.77**
IResNet50 *MagFace*	E	26.8	**9.2**	3.5	**3.2**	3.3	**9.9**	**1.6**	0.8	0.5	0.4	28.0	**11.0**	5.9	5.2	5.1	**9.0**	3.4	2.3	1.9	1.9
	C	**1.**	**.91**	.76	.74	**.75**	.88	.61	**.80**	.81	.83	1.	**.91**	.72	.69	**.66**	**.85**	**.64**	.52	.51	.52
	κ	**.97**	**.74**	.52	.51	**.52**	.88	.58	**.76**	**.77**	.78	.97	**.87**	.69	.67	**.65**	**.90**	**.72**	.57	.58	.57
	L	**1.**	**.96**	.88	.86	**.87**	.84	.62	**.81**	.82	.84	.99	**.88**	.70	.68	**.65**	**.76**	**.54**	.61	.60	.61
	S	.98	**.81**	**.63**	**.54**	**.56**	.95	.61	.81	**.82**	**.81**	.96	**.82**	**.66**	**.60**	**.63**	.92	.76	.68	.64	.68
IResNet100 *ArcFace*	E	**26.6**	**9.1**	2.9	**2.4**	2.5	**9.4**	**1.5**	0.6	**0.3**	0.4	**27.2**	**11.4**	5.1	4.4	**4.2**	**9.0**	**3.1**	**2.0**	**1.8**	1.9
	C	.99	**.90**	.77	.73	**.74**	**.87**	.60	.76	**.81**	.82	1.	**.91**	**.75**	.72	.65	**.84**	**.63**	.50	.50	.50
	κ	.96	**.73**	.54	.51	.52	**.88**	.58	.73	**.77**	.77	**.97**	**.87**	**.72**	.69	.64	.89	**.71**	.58	**.59**	.59
	L	1.	**.95**	.88	.86	**.86**	**.84**	.61	.77	**.82**	.83	.99	**.88**	**.74**	.71	.63	**.75**	**.53**	.59	.58	.59
	S	**.98**	.79	.55	.52	.51	.94	.65	.79	.82	.79	**.96**	.81	.63	.54	.61	**.93**	.77	.72	.70	.75
IResNet100 *CosFace*	E	29.0	9.3	3.1	2.6	2.3	11.9	1.7	**0.5**	0.4	**0.3**	29.6	11.8	**5.1**	**4.2**	4.5	9.7	3.5	2.2	1.9	**1.8**
	C	.99	.89	.76	.71	.73	.80	**.63**	**.80**	.80	.82	1.	.90	.73	**.74**	**.70**	.81	.60	**.52**	.52	.50
	κ	.92	.67	.50	**.54**	.50	.80	**.60**	**.76**	.76	.77	.95	.86	.70	**.71**	**.68**	.87	.68	.57	.57	**.59**
	L	.99	.95	.88	.84	.86	.76	**.64**	**.81**	.82	.83	.99	.87	.71	**.72**	**.68**	.72	.50	**.61**	.61	.58
	S	.95	.80	**.58**	.54	.50	**.94**	.62	.78	.82	.79	.94	.79	.60	.52	.58	.93	**.80**	**.77**	**.74**	**.77**
IResNet100 *MagFace*	E	27.3	9.4	**2.6**	2.5	**2.1**	10.1	1.8	0.6	0.3	0.5	27.6	11.5	5.5	4.9	4.5	9.3	3.5	2.0	1.9	1.9
	C	**.99**	.88	**.79**	**.74**	.68	.83	.62	.77	.81	**.83**	**1.**	.90	.73	.70	.63	.84	.61	.52	**.53**	**.51**
	κ	**.96**	.67	**.56**	.51	**.54**	.84	.59	.74	.77	**.78**	.97	.86	.71	.67	.63	**.89**	.70	**.60**	.56	.58
	L	**1.**	.94	**.90**	**.86**	.83	.80	.63	.78	.82	**.84**	**.99**	.87	.72	.69	.63	.74	.51	.57	**.61**	**.61**
	S	.97	**.81**	.55	**.56**	**.60**	.94	**.65**	**.82**	**.82**	**.81**	.96	**.81**	**.65**	**.62**	**.66**	.91	.75	.67	.67	.72

3.3. Epoch 20

Table 3. **Verification performance and membership separability at epoch 20.** Column conventions as described in Section 3. **Bold**: best EER or AUC per backbone.

Backbone Head		LFW 1K	10K	50K	100K	All	XQLFW 1K	10K	50K	100K	All	CPLFW 1K	10K	50K	100K	All	RFW 1K	10K	50K	100K	All
IResNet34 *ArcFace*	E	**8.0**	1.6	0.5	**0.4**	0.5	**22.0**	14.4	9.3	9.7	**10.0**	**39.1**	**21.3**	13.4	13.0	**11.6**	34.7	**18.1**	9.4	8.0	6.6
	C	.98	.57	**.76**	.78	**.79**	.98	**.95**	.82	.79	**.83**	1.	.95	.77	**.74**	.74	1.	.85	.53	.52	.52
	κ	**.76**	.76	.91	.91	**.91**	**.79**	**.62**	.58	.59	.57	**.95**	.56	.77	.78	.78	**.97**	**.57**	.72	.75	.74
	L	**1.**	.81	.52	**.51**	**.54**	1.	**.99**	.93	.92	**.94**	**1.**	**1.**	.96	.94	.95	1.	.93	.74	.70	.70
	S	.75	**.78**	**.92**	**.93**	**.92**	.92	.72	.53	.51	.53	.79	.59	.81	.82	.82	.98	.73	**.64**	**.70**	**.69**
IResNet34 *CosFace*	E	9.4	1.6	**0.4**	0.4	0.4	22.9	**14.0**	**9.2**	**9.0**	10.6	39.4	22.5	**12.9**	**12.8**	12.3	**34.2**	18.5	**9.3**	8.3	6.6
	C	.97	**.58**	.76	.77	.79	.97	.94	.80	.78	.81	1.	**.95**	**.78**	.74	**.75**	1.	**.86**	**.54**	.51	.52
	κ	.60	**.78**	**.91**	**.91**	.91	.61	.53	**.62**	**.63**	**.61**	.79	**.64**	**.80**	**.82**	**.80**	.93	.54	**.73**	**.76**	**.75**
	L	.99	**.83**	.52	.51	.53	.99	.99	.92	.91	.93	1.	.99	**.96**	.94	.95	1.	**.94**	**.74**	**.70**	.69
	S	.64	.78	.92	.92	.91	.92	.73	.59	.57	.60	.52	**.65**	**.82**	**.82**	**.84**	.96	.76	.60	.66	.67
IResNet34 *MagFace*	E	8.8	**1.3**	0.4	0.5	**0.3**	22.9	14.9	9.8	9.1	10.1	39.5	23.3	13.9	13.0	12.0	34.6	19.4	9.4	**8.0**	**6.4**
	C	**.99**	.55	.75	**.78**	.79	**.99**	.95	**.83**	**.81**	.82	**1.**	.94	.77	.74	.75	**1.**	.82	.53	**.52**	**.52**
	κ	.75	.77	.90	.91	.91	.78	.61	.56	.58	.58	.95	.56	.77	.78	.78	.97	.54	.73	.75	.75
	L	1.	.80	**.53**	.51	.52	**1.**	.99	**.94**	**.93**	.94	1.	.99	.96	**.95**	**.95**	**1.**	.92	.74	.70	**.70**
	S	**.77**	.74	.91	.92	.92	**.93**	**.75**	**.60**	**.59**	**.61**	**.85**	.53	.74	.76	.75	**.99**	**.77**	.57	.63	.62
IResNet50 *ArcFace*	E	**8.2**	1.5	**0.4**	0.5	**0.3**	21.4	**14.3**	9.9	8.4	10.0	39.7	22.1	**12.8**	11.9	11.3	34.4	19.5	8.7	**6.8**	5.9
	C	.98	.58	.75	.78	.79	**.98**	.95	.82	.81	.81	**1.**	.95	**.78**	**.75**	**.75**	**1.**	.84	.54	.50	**.52**
	κ	**.77**	.74	.90	.91	.91	**.82**	**.67**	.56	.58	.58	**.96**	.53	.76	.78	.77	**.97**	.57	.72	.74	.75
	L	**1.**	.81	**.54**	.51	.53	**1.**	.99	.94	.93	.93	**1.**	.99	.96	.95	.95	**1.**	.92	.75	**.71**	.69
	S	.77	.75	**.92**	**.92**	**.92**	.93	.71	.55	.51	.58	.83	.53	.81	**.85**	**.82**	.99	.77	**.65**	**.71**	**.68**
IResNet50 *CosFace*	E	9.6	1.7	0.5	0.4	**0.3**	22.3	14.8	9.6	8.4	9.2	39.4	22.0	12.8	**11.7**	**11.0**	34.8	**18.3**	**8.7**	6.9	**5.7**
	C	.97	.61	**.75**	.78	**.80**	.97	.95	.80	.76	.81	1.	.96	.77	.74	.75	1.	**.87**	**.56**	.50	.51
	κ	.62	**.76**	**.91**	**.91**	**.92**	.62	.56	**.61**	**.64**	**.60**	.83	**.61**	**.81**	**.81**	**.81**	.94	.56	**.72**	**.75**	**.75**
	L	.99	.84	.52	**.51**	**.54**	.99	.99	.92	.90	.93	1.	1.	.95	.94	.95	1.	**.94**	**.75**	.71	.70
	S	.66	**.76**	.92	.92	.91	**.93**	.77	.60	.57	.58	.55	**.65**	**.83**	.84	.81	.97	.79	.61	.67	.66
IResNet50 *MagFace*	E	8.8	**1.4**	0.6	**0.4**	0.4	**21.3**	15.0	**9.3**	**8.4**	**9.1**	**39.4**	**21.8**	13.1	11.9	11.2	**34.3**	19.3	9.1	7.0	6.1
	C	**.98**	**.62**	.75	**.78**	.79	.98	**.96**	**.85**	**.82**	**.82**	1.	**.96**	.78	.74	.75	1.	.86	.54	**.51**	.51
	κ	.76	.73	.90	.91	.91	.79	.66	.53	.56	.57	.95	.51	.76	.78	.78	.97	**.60**	.72	.74	.74
	L	1.	**.84**	.53	.50	.52	1.	**.99**	**.95**	**.94**	**.93**	1.	**1.**	**.96**	**.95**	**.95**	1.	.94	.74	.71	**.70**
	S	**.79**	.71	.91	.92	.91	.93	**.80**	**.60**	**.57**	**.61**	**.87**	.56	.76	.76	.76	**.99**	**.81**	.57	.63	.62
IResNet100 *ArcFace*	E	**8.6**	**1.3**	**0.4**	**0.3**	0.4	**20.9**	**14.1**	9.4	**8.1**	8.6	**38.5**	**20.8**	**11.2**	**10.6**	10.4	34.3	**18.0**	**7.7**	**5.4**	**4.6**
	C	.98	.60	**.74**	.78	**.80**	**.99**	.96	**.86**	**.85**	**.84**	1.	.95	.78	**.75**	.74	1.	**.87**	.56	.51	.51
	κ	**.78**	.73	.90	.91	**.91**	**.83**	**.67**	.51	.53	.53	**.96**	.52	.74	.76	.77	**.98**	**.61**	.70	.73	.74
	L	**1.**	.83	.54	.51	**.54**	**1.**	.99	**.95**	**.94**	**.94**	1.	.99	.96	**.95**	.95	**1.**	**.94**	.75	.71	.69
	S	.78	.73	.92	**.92**	**.92**	.93	.77	.56	.53	.58	.84	.52	.82	**.84**	.82	.99	.79	**.63**	**.71**	**.69**
IResNet100 *CosFace*	E	9.0	1.5	0.5	**0.3**	0.4	21.9	14.4	9.2	8.2	9.2	38.7	21.1	11.5	10.8	10.2	34.4	19.1	8.0	5.7	4.7
	C	.97	.60	.74	.77	.78	.97	.95	.84	.81	.84	1.	.95	**.78**	.74	.74	1.	.85	**.56**	**.51**	.51
	κ	.65	**.76**	**.90**	.90	.91	.66	.59	**.57**	**.59**	**.56**	.87	**.59**	**.79**	**.79**	**.80**	.94	.55	**.72**	.74	**.75**
	L	1.	.83	**.54**	.51	.53	.99	.99	.94	.93	.94	1.	.99	**.96**	.95	.94	1.	.93	**.76**	**.72**	.69
	S	.68	**.75**	**.92**	.92	.91	**.94**	.75	.60	.58	.61	.61	**.61**	**.84**	.83	**.83**	.97	.80	.59	.68	.67
IResNet100 *MagFace*	E	8.9	1.5	0.5	0.4	**0.3**	22.1	15.6	**8.9**	9.3	**7.8**	39.8	21.3	11.9	10.6	**10.2**	**34.0**	18.8	8.3	5.9	4.7
	C	**.98**	**.62**	.74	**.78**	.79	.98	**.96**	.85	.82	.83	**1.**	**.96**	.77	.73	**.74**	**1.**	.86	.56	.51	**.52**
	κ	.77	.72	.90	**.91**	.91	.82	.67	.52	.55	.54	.96	.53	.74	.78	.77	.97	.61	.70	**.74**	.74
	L	1.	**.84**	.54	**.51**	.53	1.	**.99**	.95	.93	.94	**1.**	**1.**	.96	.94	**.95**	1.	.94	.75	.70	**.70**
	S	**.80**	.71	.91	.92	.91	.93	**.81**	**.61**	**.62**	**.61**	**.88**	.58	.75	.78	.77	**.99**	**.81**	.56	.63	.63

Backbone Head		CFP-FP 1K	10K	50K	100K	All	CFP-FF 1K	10K	50K	100K	All	AgeDB 1K	10K	50K	100K	All	IJB-C 1K	10K	50K	100K	All
IResNet34 *ArcFace*	E	**26.3**	**8.7**	3.4	2.9	2.9	**10.0**	1.6	**0.4**	0.6	**0.3**	27.3	10.7	5.3	4.9	**4.1**	**8.3**	**3.5**	2.4	2.0	1.7
	C	**1.**	.98	.74	**.73**	.72	1.	**.60**	**.79**	**.81**	**.82**	**1.**	**.98**	.73	.69	**.68**	**.98**	**.75**	**.52**	**.52**	.50
	κ	**1.**	**.89**	.52	.50	.50	1.	**.61**	**.76**	**.77**	**.78**	**1.**	**.95**	**.71**	.66	**.66**	**.99**	**.81**	.57	.56	.59
	L	**1.**	.99	.86	**.86**	.85	**1.**	**.57**	**.80**	**.82**	**.84**	**1.**	**.96**	**.72**	.68	**.66**	**.96**	**.66**	**.61**	**.61**	.59
	S	**1.**	.90	.58	.52	.51	**1.**	.53	.78	.81	.79	**.99**	.91	.64	.57	.59	**1.**	.90	.76	.72	.74
IResNet34 *CosFace*	E	27.7	9.4	**3.1**	2.9	2.9	10.2	**1.6**	0.6	**0.5**	0.4	**27.2**	11.1	5.4	4.7	4.7	8.9	3.8	2.2	2.0	1.8
	C	1.	**.98**	**.78**	.70	.72	1.	.56	.78	.80	.82	**1.**	.98	**.74**	**.72**	.67	.96	.74	.51	.52	.51
	κ	1.	.88	.54	**.55**	**.53**	1.	.57	.74	.76	.77	.99	.94	.71	**.69**	.66	.98	.81	.58	.57	.58
	L	1.	**.99**	**.89**	.84	.85	.99	.54	.79	.81	.84	1.	.96	.72	**.71**	.65	.92	.65	.60	.60	**.60**
	S	.99	.90	.55	.50	.51	.99	**.56**	.79	.81	.77	.98	.91	.64	.54	.58	.99	**.90**	**.78**	**.75**	**.76**
IResNet34 *MagFace*	E	27.7	10.2	3.7	**2.8**	**2.7**	**10.0**	1.9	0.6	0.5	0.5	27.5	**10.4**	**5.3**	**4.7**	4.4	8.6	4.0	**2.0**	**1.8**	**1.7**
	C	**1.**	.97	.77	.72	**.73**	**1.**	.53	.79	.81	.82	**1.**	.97	.71	.70	.65	.98	.75	.51	.51	**.52**
	κ	**1.**	.87	**.54**	.50	.50	**1.**	.55	.75	.77	.77	1.	.94	.69	.67	.64	.99	.81	**.60**	**.59**	**.61**
	L	1.	.98	.88	.85	**.86**	1.	.51	.80	.81	.83	**1.**	.95	.70	.69	.64	.95	.65	.58	.58	.57
	S	1.	**.93**	**.60**	**.53**	**.58**	1.	.53	**.80**	**.83**	**.81**	.99	**.91**	**.69**	**.60**	**.64**	.99	.86	.69	.66	.68
IResNet50 *ArcFace*	E	**26.6**	**9.1**	**2.6**	2.6	2.7	10.2	1.6	**0.5**	**0.3**	0.2	**26.5**	**10.4**	**5.1**	**3.9**	4.2	**8.3**	**3.4**	2.2	1.7	**1.5**
	C	**1.**	.96	**.79**	.73	**.75**	**1.**	.61	**.77**	.80	.81	**1.**	.98	.75	.72	.68	**.98**	.77	.50	**.52**	**.52**
	κ	**1.**	.88	**.57**	.51	**.52**	**1.**	.63	**.74**	.76	.76	1.	.94	.72	.69	.67	.99	.83	.59	**.60**	**.60**
	L	**1.**	.98	.89	.86	**.87**	**1.**	.58	**.78**	.80	.82	**1.**	.96	.73	.71	.67	**.96**	.68	**.59**	.57	.57
	S	1.	.92	.54	.52	.50	1.	.56	.79	.80	.79	.99	.92	.64	.56	.58	**1.**	.90	.76	.71	.74
IResNet50 *CosFace*	E	28.4	10.0	3.1	2.4	**2.1**	**10.2**	**1.5**	0.6	0.4	**0.2**	27.8	11.8	5.2	4.4	**4.0**	8.8	3.9	2.1	1.8	1.7
	C	1.	**.98**	.75	**.73**	.73	1.	.60	.77	.79	.82	**1.**	**.99**	**.77**	**.74**	.68	.96	.75	.50	.50	.51
	κ	1.	.89	.50	**.51**	.52	1.	.61	.73	.75	.77	.99	.95	**.74**	**.70**	.66	.98	.81	.59	.58	.59
	L	1.	**.99**	.87	**.86**	.86	.99	.58	.78	.80	**.83**	1.	**.97**	**.75**	**.72**	.66	.93	.66	.59	**.58**	**.60**
	S	.99	.92	.54	.55	.51	1.	.59	.77	.80	.78	.98	.91	.62	.55	.60	.99	**.92**	**.78**	**.73**	**.77**
IResNet50 *MagFace*	E	26.9	9.5	2.9	**2.3**	2.4	10.2	1.9	0.6	**0.3**	0.3	28.1	10.8	5.2	4.2	4.3	8.6	3.4	**1.9**	**1.7**	1.7
	C	1.	.98	.79	.73	.72	1.	**.63**	.77	**.80**	**.82**	1.	.98	.76	.71	**.68**	.98	**.78**	**.55**	.51	.50
	κ	1.	**.91**	.57	.50	.50	1.	**.65**	.74	**.76**	**.78**	**1.**	**.95**	.73	.69	**.67**	**.99**	**.84**	**.63**	.59	.59
	L	1.	.99	**.89**	.86	.86	1.	**.60**	.78	**.81**	.83	1.	.97	.74	.70	**.67**	.96	**.69**	.54	.58	.59
	S	**1.**	**.93**	**.57**	**.56**	**.58**	**1.**	**.60**	**.79**	**.83**	**.80**	**.99**	**.92**	**.68**	**.62**	**.64**	.99	.88	.69	.69	.71
IResNet100 *ArcFace*	E	**26.7**	**8.2**	**2.1**	**1.8**	1.8	9.9	**1.4**	**0.4**	0.3	0.3	26.9	**10.8**	**4.5**	**3.6**	3.8	**8.2**	**3.1**	**1.7**	1.5	1.5
	C	**1.**	.97	.80	.75	.73	**1.**	**.64**	.75	.79	**.82**	**1.**	**.99**	.79	.72	**.71**	**.98**	.78	.55	**.54**	.51
	κ	**1.**	.91	.59	**.53**	.51	**1.**	**.66**	.72	.75	**.77**	**1.**	**.96**	.76	.69	**.69**	**.99**	.84	**.63**	**.62**	.60
	L	**1.**	.99	.90	.87	.86	**1.**	**.62**	.76	.80	**.83**	**1.**	**.97**	.77	.71	**.69**	**.96**	.70	.54	.54	.58
	S	1.	.92	.52	.56	.52	**1.**	**.62**	.76	.81	.78	**.99**	.93	.65	.56	.60	**1.**	**.92**	.76	.73	.76
IResNet100 *CosFace*	E	28.7	9.2	2.3	2.0	1.7	10.9	1.7	0.4	0.4	0.3	26.9	11.9	4.7	3.7	**3.5**	8.6	3.8	1.8	**1.5**	1.6
	C	**1.**	.98	.78	**.75**	.73	1.	.61	**.76**	.78	.81	**1.**	.99	**.80**	**.74**	.70	.97	.75	.53	.52	.50
	κ	1.	.88	.54	.51	.50	1.	.63	**.72**	.74	.76	.99	.95	**.77**	**.71**	.68	.98	.81	.62	.60	.59
	L	1.	.99	.89	**.87**	.86	.99	.58	**.77**	.79	.82	1.	.96	**.78**	**.73**	.68	.93	.67	**.55**	.56	**.58**
	S	1.	.91	.51	**.60**	**.56**	1.	.58	.77	.80	.77	.98	.91	.65	.56	.57	.99	.91	**.78**	**.75**	**.79**
IResNet100 *MagFace*	E	26.8	8.4	2.3	1.9	**1.7**	**9.8**	1.8	0.4	**0.3**	**0.2**	**26.7**	10.8	4.9	3.8	3.7	8.3	3.4	1.8	1.6	**1.4**
	C	**1.**	**.98**	**.81**	.73	**.74**	1.	.64	.74	**.80**	.81	**1.**	.98	.79	.71	.68	.98	**.79**	**.55**	.52	**.54**
	κ	**1.**	**.93**	**.61**	.51	**.52**	1.	.65	.71	**.76**	.77	1.	.96	.76	.68	.67	.99	**.84**	.63	.60	**.62**
	L	1.	**.99**	**.90**	.85	**.87**	1.	.61	.75	**.80**	.83	**1.**	.97	.77	.70	.67	.95	**.70**	.54	**.57**	.56
	S	**1.**	**.94**	**.56**	.51	.54	1.	.58	**.79**	**.82**	**.80**	.99	**.93**	**.70**	**.62**	**.64**	.99	.88	.70	.69	.70

3.4. Epoch 40

Table 4. **Verification performance and membership separability at epoch 40.** Column conventions as described in Section 3. **Bold**: best EER or AUC per backbone.

Backbone Head		LFW					XQLFW				
		1K	10K	50K	100K	All	1K	10K	50K	100K	All
IResNet34 *ArcFace*	*E*	**7.7**	1.3	0.4	**0.2**	0.3	**21.3**	**12.8**	**7.3**	5.8	5.3
	C	.99	**.92**	.56	.71	**.77**	.99	**1.**	**.96**	.91	.89
	κ	**.78**	**.60**	.84	.89	.90	**.83**	**.94**	**.70**	.58	.54
	L	1.	**.98**	**.74**	**.59**	**.51**	1.	**1.**	.99	.97	.96
	S	.78	.59	.86	.90	.90	.93	.93	.63	.54	.55
IResNet34 *CosFace*	*E*	9.3	1.3	**0.4**	0.2	0.3	22.8	13.3	7.5	**5.7**	5.5
	C	.97	.89	**.58**	.71	.77	.97	.99	.94	.91	.88
	κ	.62	.51	**.86**	**.89**	**.91**	.65	.88	.61	.53	.50
	L	.99	.97	.71	.59	.51	1.	1.	.99	.97	.96
	S	.67	.53	**.87**	**.91**	.90	.93	.90	.62	.54	.56
IResNet34 *MagFace*	*E*	8.9	**1.2**	0.5	0.4	**0.2**	23.0	13.8	7.5	5.9	**5.1**
	C	**.99**	.91	.58	**.72**	.77	**.99**	1.	.96	**.91**	**.89**
	κ	.78	.57	.84	.89	.90	.80	.93	.69	**.59**	**.54**
	L	**1.**	.98	.72	.59	.51	**1.**	1.	**.99**	**.97**	**.96**
	S	**.79**	**.60**	.86	.90	**.90**	**.93**	**.95**	**.72**	**.63**	**.64**
IResNet50 *ArcFace*	*E*	**8.2**	1.2	0.4	**0.2**	0.2	**21.0**	**12.5**	7.0	**5.2**	4.5
	C	.99	**.92**	.50	.69	.76	.99	**1.**	**.96**	**.92**	.89
	κ	.80	**.62**	.81	.88	.90	**.86**	**.95**	**.74**	**.62**	.57
	L	1.	**.98**	**.77**	**.62**	**.53**	1.	1.	**.99**	**.98**	.97
	S	.80	.61	.84	.89	.89	**.94**	.94	.67	.57	.56
IResNet50 *CosFace*	*E*	9.4	1.5	0.3	0.2	**0.1**	22.0	13.9	7.3	5.5	4.9
	C	.98	.89	**.54**	.69	.76	.97	.99	.96	.92	.89
	κ	.65	.51	**.84**	**.89**	**.90**	.66	.89	.65	.56	.52
	L	1.	.97	.75	.61	.52	.99	**1.**	.99	.97	.96
	S	.69	.55	**.85**	**.90**	.89	.93	.91	.66	.56	.58
IResNet50 *MagFace*	*E*	8.6	**1.1**	**0.3**	0.3	0.1	21.3	13.4	**7.0**	5.3	**4.3**
	C	**.99**	.92	.53	**.70**	**.76**	**.99**	1.	.96	.92	**.89**
	κ	**.80**	.59	.82	.88	.90	.81	.93	.72	.62	**.57**
	L	**1.**	.98	.76	.61	.52	**1.**	1.	.99	.97	**.97**
	S	**.81**	**.62**	.84	.89	**.90**	.93	**.95**	**.75**	**.65**	**.64**
IResNet100 *ArcFace*	*E*	**8.5**	1.2	**0.3**	0.2	0.3	**20.2**	**12.8**	**6.1**	**5.1**	**4.2**
	C	.99	**.92**	**.54**	.65	**.75**	**.99**	.99	**.97**	**.94**	**.91**
	κ	**.83**	**.63**	.78	.87	.90	**.88**	**.96**	**.78**	**.68**	**.61**
	L	**1.**	**.98**	**.80**	**.67**	**.55**	1.	1.	**.99**	**.98**	**.97**
	S	.82	**.62**	.82	.89	.88	.94	.94	.72	.61	.58
IResNet100 *CosFace*	*E*	8.7	1.1	0.3	0.3	0.1	21.8	13.5	7.2	5.2	4.2
	C	.98	.89	.51	**.66**	.75	.97	.99	.97	.93	.90
	κ	.69	.53	**.82**	**.88**	**.90**	.71	.90	.71	.62	.57
	L	1.	.97	.77	.65	.54	1.	1.	.99	.98	.97
	S	.72	.56	**.84**	**.89**	.88	**.95**	.92	.72	.60	.60
IResNet100 *MagFace*	*E*	8.8	**1.1**	**0.3**	**0.1**	**0.1**	22.0	12.9	6.4	**5.1**	4.3
	C	**.99**	.91	.51	.65	.75	.99	**1.**	.97	.94	.91
	κ	.81	.58	.81	.87	.90	.86	.94	.76	.67	.60
	L	1.	.98	.78	.66	.54	**1.**	**1.**	.99	.98	.97
	S	**.83**	.61	.82	.88	**.89**	.94	**.95**	**.78**	**.69**	**.65**

Backbone Head		CPLFW					RFW				
		1K	10K	50K	100K	All	1K	10K	50K	100K	All
IResNet34 *ArcFace*	*E*	39.6	20.7	10.7	9.0	8.4	34.6	17.4	6.3	**3.9**	2.8
	C	**1.**	**1.**	**.90**	.80	.75	1.	**.99**	**.78**	.61	.52
	κ	**.96**	**.90**	.61	.71	.74	**.98**	**.93**	.52	.67	.72
	L	**1.**	**1.**	**.99**	**.97**	.95	1.	1.	**.90**	.79	.73
	S	.83	.82	.78	.83	.83	**.99**	.97	.56	**.63**	**.66**
IResNet34 *CosFace*	*E*	**39.4**	21.0	**10.6**	9.0	**8.1**	**34.2**	17.6	**6.1**	4.0	**2.6**
	C	1.	1.	.89	.80	**.75**	1.	.98	.77	**.61**	**.53**
	κ	.83	.80	**.69**	**.76**	**.78**	.94	.89	**.57**	**.69**	**.72**
	L	1.	1.	.99	.97	**.95**	1.	.99	.89	**.80**	**.73**
	S	.57	.67	**.83**	**.86**	**.84**	.97	.95	.56	.62	.65
IResNet34 *MagFace*	*E*	40.1	**20.6**	10.7	**8.9**	8.3	34.3	**17.4**	6.2	4.0	2.8
	C	1.	1.	.89	**.81**	.75	**1.**	.99	.77	.60	.52
	κ	.95	.88	.62	.71	.74	.98	.93	.53	.67	.72
	L	1.	1.	.99	.97	.95	**1.**	**1.**	.89	.79	.72
	S	**.88**	**.86**	.70	.78	.79	.99	**.97**	**.64**	.55	.59
IResNet50 *ArcFace*	*E*	**39.3**	20.6	10.0	**8.3**	7.7	34.3	17.5	6.0	3.5	2.1
	C	**1.**	**1.**	**.92**	.82	.75	1.	**.99**	**.83**	**.64**	.53
	κ	**.98**	**.90**	.57	.70	.74	**.98**	**.94**	**.56**	.64	.71
	L	1.	1.	**.99**	.97	.95	1.	**1.**	**.92**	**.82**	.74
	S	.86	.83	.76	.83	**.83**	.99	.97	.62	**.60**	**.66**
IResNet50 *CosFace*	*E*	39.5	**20.2**	**9.6**	8.6	7.7	34.9	17.2	5.9	**3.3**	**2.1**
	C	1.	1.	.91	**.82**	**.76**	1.	.99	.81	.64	**.54**
	κ	.86	.83	**.66**	**.74**	**.77**	.95	.90	.51	**.67**	**.72**
	L	1.	1.	.99	**.97**	.95	1.	.99	.91	.81	**.74**
	S	.60	.70	**.80**	**.85**	.83	.97	.96	.62	.60	.63
IResNet50 *MagFace*	*E*	39.6	20.6	9.8	8.5	**7.6**	**34.0**	**17.2**	**5.8**	3.5	2.3
	C	1.	1.	.91	.82	.75	**1.**	.99	.81	.63	.53
	κ	.96	.88	.59	.70	.74	.98	.93	.52	.65	.71
	L	**1.**	**1.**	.99	.97	**.96**	**1.**	1.	.91	.80	.73
	S	**.90**	**.87**	.69	.78	.79	**.99**	**.97**	**.68**	.53	.59
IResNet100 *ArcFace*	*E*	**38.0**	**19.3**	9.1	7.7	7.0	34.2	16.8	5.2	2.8	1.5
	C	**1.**	1.	**.93**	**.84**	.76	1.	**.99**	**.86**	**.69**	.55
	κ	**.98**	**.91**	.52	.67	.73	**.98**	**.95**	**.61**	.60	.69
	L	**1.**	1.	**.99**	**.98**	**.96**	1.	**1.**	**.93**	**.84**	.75
	S	.88	.84	.72	.81	**.82**	**.99**	.97	.66	**.57**	**.64**
IResNet100 *CosFace*	*E*	38.7	19.6	9.1	**7.5**	7.0	34.2	16.4	**5.0**	**2.5**	1.6
	C	1.	1.	.92	.84	**.76**	1.	.99	.83	.68	**.56**
	κ	.90	.84	**.62**	**.72**	**.76**	.96	.91	.54	**.63**	**.71**
	L	1.	1.	.99	.97	.96	1.	.99	.92	.84	**.75**
	S	.66	.72	**.77**	**.83**	.82	.98	.96	.66	.56	.61
IResNet100 *MagFace*	*E*	40.0	20.3	**8.9**	7.9	**7.0**	**33.8**	**16.3**	5.0	3.0	**1.4**
	C	1.	**1.**	.92	.84	.76	**1.**	.99	.83	.68	.55
	κ	.96	.89	.55	.67	.73	.98	.93	.57	.61	.70
	L	1.	**1.**	.99	.98	.96	**1.**	1.	.92	.83	.74
	S	**.91**	**.87**	.67	.77	.79	.99	**.97**	**.71**	.51	.59

Backbone Head		CFP-FP					CFP-FF				
		1K	10K	50K	100K	All	1K	10K	50K	100K	All
IResNet34 *ArcFace*	*E*	**26.3**	**7.7**	**1.8**	**1.1**	1.0	**9.6**	1.4	0.3	0.3	**0.1**
	C	**1.**	**1.**	**.94**	.83	.76	**1.**	**.99**	.53	.71	.78
	κ	**1.**	**1.**	**.82**	.63	.55	**1.**	**.99**	.51	.68	.73
	L	**1.**	**1.**	**.97**	.91	.87	**1.**	**.96**	.54	.72	.79
	S	1.	**1.**	.65	.58	.61	1.	**.98**	.55	.72	.73
IResNet34 *CosFace*	*E*	27.8	8.7	1.8	1.2	0.9	10.2	**1.2**	**0.3**	0.3	0.1
	C	**1.**	1.	.93	.83	.75	1.	.96	**.58**	**.72**	.78
	κ	1.	.99	.77	.59	.52	1.	.96	**.55**	.69	.73
	L	1.	1.	.96	.91	.87	.99	.92	**.59**	**.73**	**.79**
	S	.99	.99	.57	**.62**	**.62**	1.	.96	.59	.73	.73
IResNet34 *MagFace*	*E*	28.5	7.9	1.9	1.2	**0.9**	10.5	1.4	0.3	**0.2**	0.2
	C	**1.**	1.	.94	**.84**	**.76**	1.	.98	.56	.72	**.78**
	κ	1.	1.	.82	**.63**	**.55**	**1.**	.98	.53	**.69**	**.74**
	L	1.	1.	.97	**.92**	**.87**	1.	.95	.57	.73	.79
	S	**1.**	1.	**.72**	.50	.53	**1.**	.97	**.62**	**.76**	**.75**
IResNet50 *ArcFace*	*E*	**26.0**	7.9	1.6	0.9	**0.6**	**9.5**	1.5	**0.3**	0.1	0.2
	C	**1.**	**1.**	**.96**	**.86**	.77	1.	**.99**	**.55**	.67	.76
	κ	**1.**	**1.**	**.88**	**.67**	.56	**1.**	**.99**	**.58**	.64	.72
	L	**1.**	**1.**	**.98**	.93	.88	1.	**.97**	**.53**	.67	.77
	S	1.	1.	.69	.57	.62	1.	**.98**	.51	.70	.72
IResNet50 *CosFace*	*E*	27.9	8.5	**1.4**	0.9	0.6	10.3	**1.3**	0.4	**0.1**	**0.1**
	C	1.	1.	.95	.85	.76	1.	.97	.51	.69	.76
	κ	1.	1.	.83	.63	.54	1.	.97	.52	.66	.71
	L	1.	1.	.97	.92	.87	.99	.94	.52	.69	.77
	S	1.	.99	.62	**.62**	**.64**	1.	.96	.52	.71	.72
IResNet50 *MagFace*	*E*	27.4	**7.2**	1.5	**0.9**	0.8	10.4	1.5	0.4	0.2	**0.1**
	C	**1.**	1.	.96	.86	**.78**	**1.**	.99	.52	**.69**	**.76**
	κ	1.	1.	.85	.66	**.57**	1.	.99	.54	**.66**	**.72**
	L	1.	1.	.98	**.93**	**.88**	**1.**	.96	.50	**.69**	**.77**
	S	**1.**	**1.**	**.75**	.52	.54	**1.**	.97	**.55**	**.73**	**.74**
IResNet100 *ArcFace*	*E*	**26.2**	**6.7**	**1.2**	**0.7**	0.6	**9.6**	1.4	0.3	**0.1**	**0.1**
	C	**1.**	**1.**	**.97**	**.89**	.78	**1.**	**.99**	**.62**	.61	.73
	κ	**1.**	**1.**	**.91**	**.72**	**.58**	**1.**	**.99**	**.65**	.59	.69
	L	**1.**	**1.**	**.98**	**.94**	**.88**	**1.**	**.97**	**.59**	.61	.75
	S	1.	1.	.71	.55	.63	1.	**.98**	**.57**	.66	.70
IResNet100 *CosFace*	*E*	28.6	7.7	1.4	0.8	**0.5**	10.9	1.4	**0.2**	0.2	**0.1**
	C	**1.**	1.	.96	.88	.77	1.	.98	.56	**.64**	.74
	κ	1.	1.	.87	.67	.55	1.	.98	.58	**.61**	.70
	L	1.	1.	.97	.94	.88	.99	.94	.54	**.64**	**.75**
	S	1.	.99	.62	**.61**	**.65**	1.	.95	.51	**.69**	.71
IResNet100 *MagFace*	*E*	27.9	7.0	1.2	**0.7**	0.7	10.1	**1.3**	0.3	0.2	0.1
	C	**1.**	1.	.96	.88	**.78**	**1.**	.99	.57	.63	**.74**
	κ	**1.**	1.	.89	.71	.58	1.	.98	.59	.61	**.70**
	L	1.	1.	.98	.94	.88	1.	.96	.54	.63	.75
	S	**1.**	**1.**	**.76**	.52	.56	**1.**	.97	.51	.68	**.71**

Backbone Head		AgeDB					IJB-C				
		1K	10K	50K	100K	All	1K	10K	50K	100K	All
IResNet34 *ArcFace*	*E*	**27.1**	**9.6**	3.6	3.1	2.8	**8.3**	2.9	1.7	1.5	1.3
	C	**1.**	1.	**.96**	.84	.75	**.99**	**.94**	**.72**	.62	**.58**
	κ	**1.**	.99	**.92**	.81	.73	**.99**	**.96**	**.78**	.69	**.65**
	L	**1.**	1.	**.93**	.81	.73	**.96**	**.90**	**.63**	.53	.52
	S	.99	.98	.85	.70	.66	**1.**	**1.**	.93	.84	.83
IResNet34 *CosFace*	*E*	27.2	10.4	3.8	3.0	**2.6**	8.7	3.1	1.7	1.5	1.4
	C	**1.**	**1.**	.96	**.85**	**.76**	.97	.92	.69	.60	.56
	κ	.99	.99	.92	**.81**	**.74**	.98	.94	.76	.68	.64
	L	1.	1.	.92	**.82**	**.73**	.93	.87	.60	.52	**.53**
	S	.98	.97	.84	.69	.65	.99	1.	**.94**	**.87**	**.85**
IResNet34 *MagFace*	*E*	27.4	10.2	**3.6**	**2.9**	2.8	8.7	**2.9**	**1.6**	**1.4**	**1.3**
	C	**1.**	1.	.96	.84	.75	.99	.94	.71	**.62**	.57
	κ	1.	**.99**	.92	.81	.73	.99	.96	.78	**.69**	.65
	L	**1.**	**1.**	.92	.81	.73	.96	.89	.62	**.53**	.52
	S	**.99**	**.99**	**.86**	**.73**	**.70**	.99	.99	.87	.79	.78
IResNet50 *ArcFace*	*E*	**26.4**	9.5	3.9	3.0	**2.5**	**8.2**	3.0	1.5	1.4	**1.2**
	C	**1.**	1.	**.98**	.88	.78	.99	**.95**	**.75**	**.65**	.59
	κ	1.	**1.**	**.95**	.85	.76	.99	**.97**	**.81**	**.71**	.66
	L	**1.**	**1.**	**.95**	.85	.75	.97	**.91**	**.67**	**.56**	.50
	S	**.99**	.99	.88	.74	.67	**1.**	**1.**	.95	.87	.85
IResNet50 *CosFace*	*E*	27.3	10.2	3.5	**2.7**	2.8	8.8	3.0	1.6	1.4	1.3
	C	**1.**	**1.**	.97	**.89**	**.79**	.97	.92	.72	.62	.57
	κ	.99	.99	.94	**.85**	**.76**	.98	.95	.78	.70	.65
	L	**1.**	1.	.95	**.85**	**.76**	.94	.88	.64	.54	**.51**
	S	.98	.97	.87	.73	.67	.99	1.	**.97**	**.89**	**.87**
IResNet50 *MagFace*	*E*	27.5	**9.4**	**3.4**	2.8	2.7	8.6	**2.8**	**1.5**	**1.3**	1.3
	C	**1.**	**1.**	.97	.87	.78	**.99**	.95	.74	.64	**.59**
	κ	**1.**	.99	.94	.84	.76	**.99**	.96	.80	.71	**.66**
	L	**1.**	1.	.94	.84	.75	**.97**	.90	.65	.55	.50
	S	.99	**.99**	**.89**	**.76**	**.72**	1.	.99	.90	.81	.80
IResNet100 *ArcFace*	*E*	26.4	9.6	3.3	**2.6**	**2.4**	**8.1**	2.8	1.5	1.3	**1.2**
	C	**1.**	**1.**	**.99**	**.92**	.82	**.99**	**.95**	**.78**	**.68**	**.61**
	κ	1.	**1.**	**.96**	**.89**	**.79**	**.99**	**.97**	**.83**	**.75**	**.68**
	L	**1.**	**1.**	**.96**	**.89**	**.78**	**.97**	**.91**	**.70**	**.60**	**.52**
	S	**.99**	.99	.90	.78	.70	**1.**	**1.**	.97	.91	.87
IResNet100 *CosFace*	*E*	26.7	9.7	3.4	2.7	2.5	8.5	2.9	1.6	1.3	1.3
	C	**1.**	**1.**	.98	.92	**.82**	.97	.92	.74	.66	.59
	κ	.99	.99	.95	.88	.79	.98	.95	.80	.73	.67
	L	1.	1.	.96	.88	.78	.94	.88	.66	.58	.51
	S	.98	.97	.88	.77	.69	.99	1.	**.98**	**.92**	**.90**
IResNet100 *MagFace*	*E*	**26.2**	**9.0**	**3.0**	2.7	**2.4**	8.5	**2.7**	**1.4**	**1.3**	1.2
	C	**1.**	1.	.98	.92	.80	.99	.94	.76	.68	.61
	κ	**1.**	.99	.95	.88	.78	.99	.96	.82	.74	.68
	L	**1.**	1.	.96	.88	.78	.97	.89	.67	.59	.52
	S	.99	**.99**	**.90**	**.81**	**.74**	.99	.99	.93	.86	.83

4. PairCos ($\bar{s}$) distributions

4.1. Vary training duration

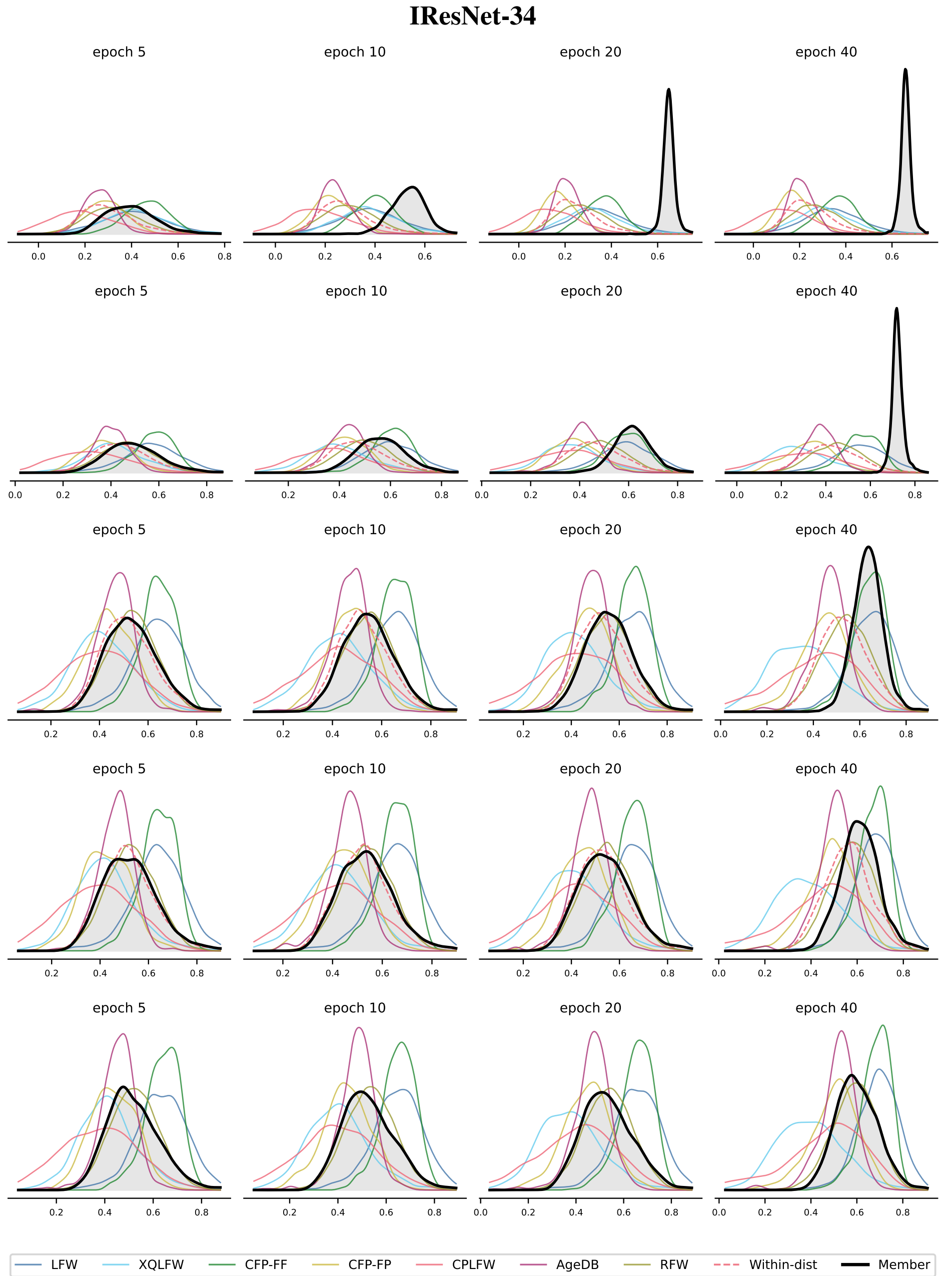


Figure 122. **Vary training duration — IResNet-34 / ArcFace.** Rows: $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

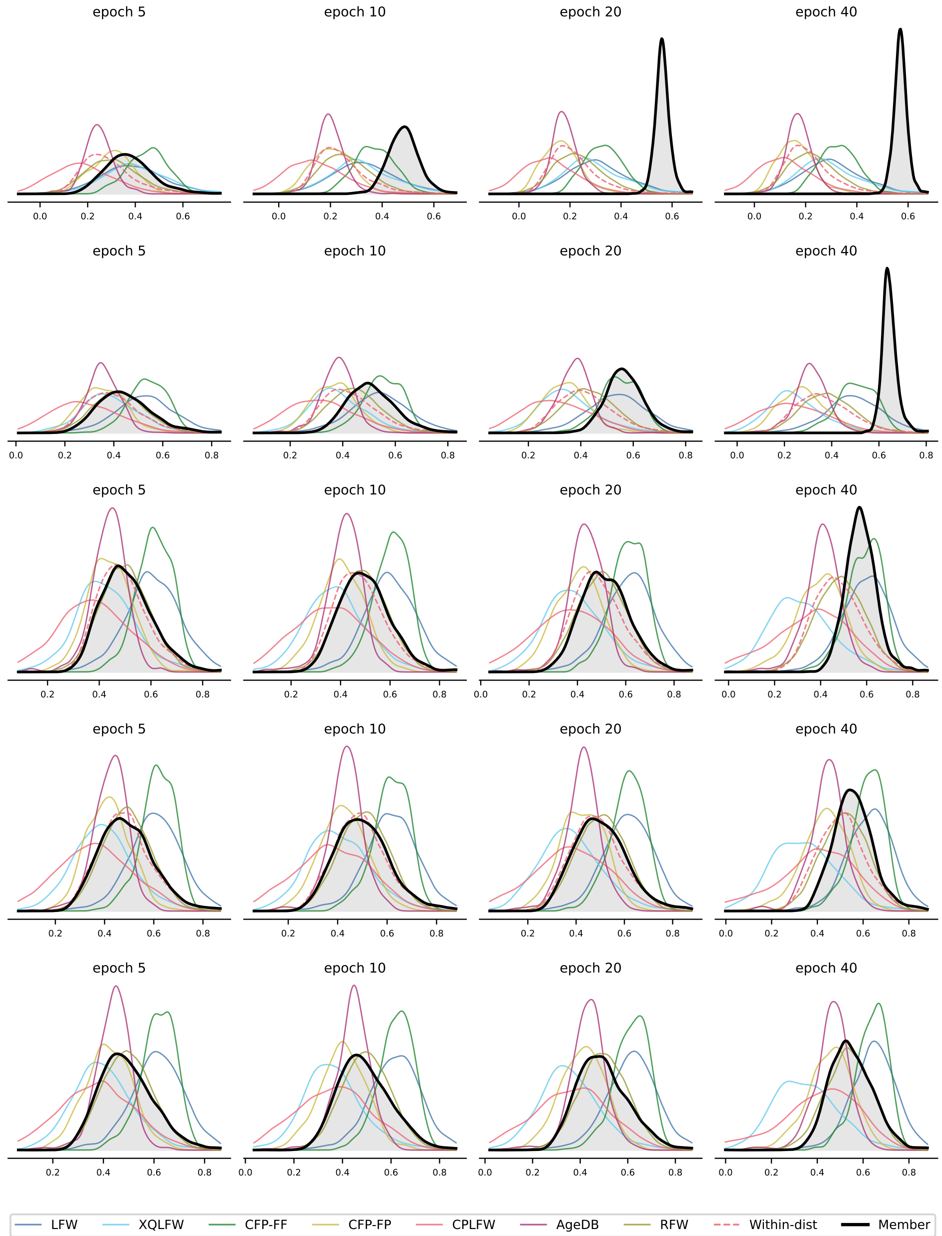


Figure 123. **Vary training duration — IResNet-34 / CosFace.** Rows: $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

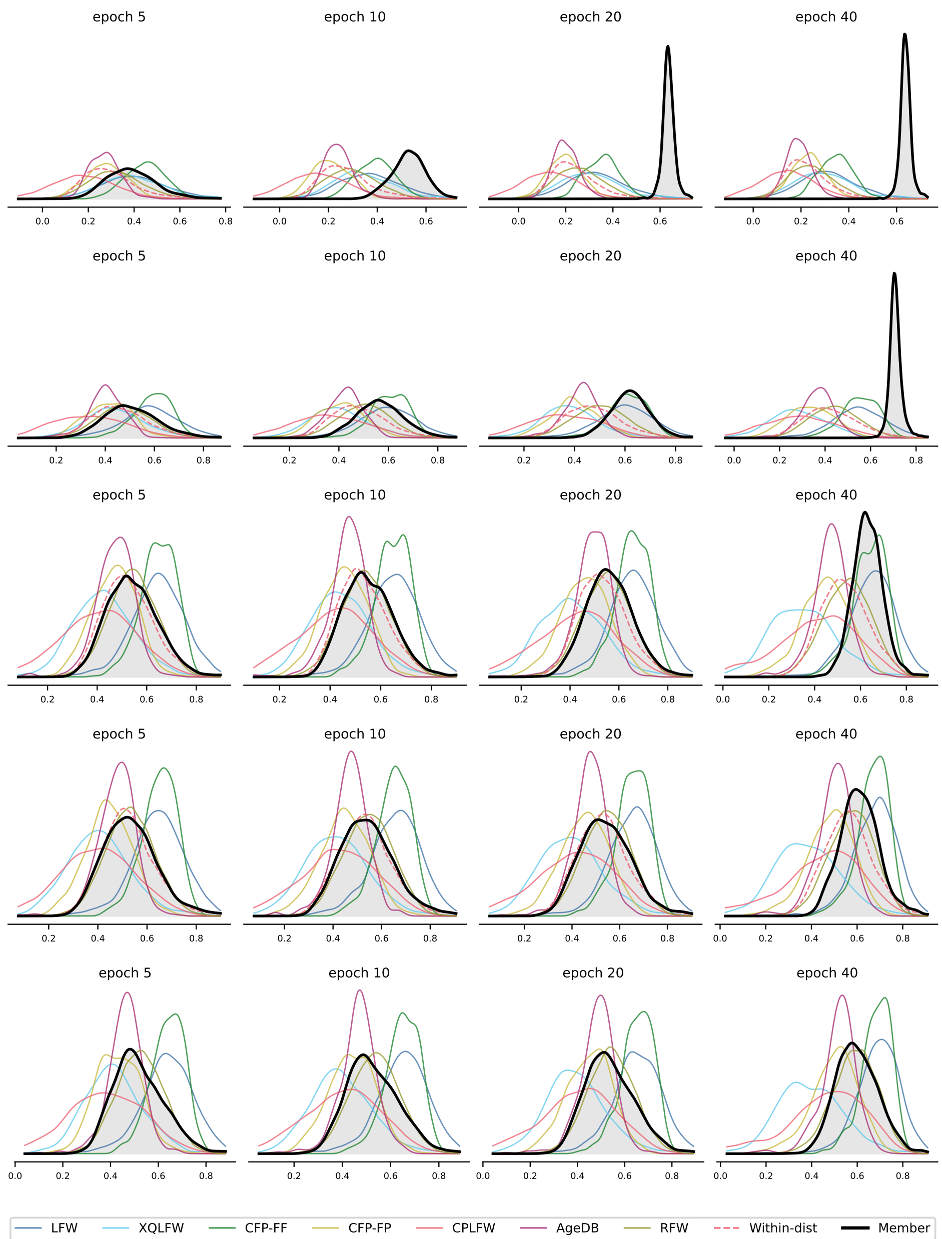


Figure 124. **Vary training duration — IResNet-34 / MagFace.** Rows: $n_{\text{ids}} \in \{1\text{K},\ 10\text{K},\ 50\text{K},\ 100\text{K},\ \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

IResNet-50

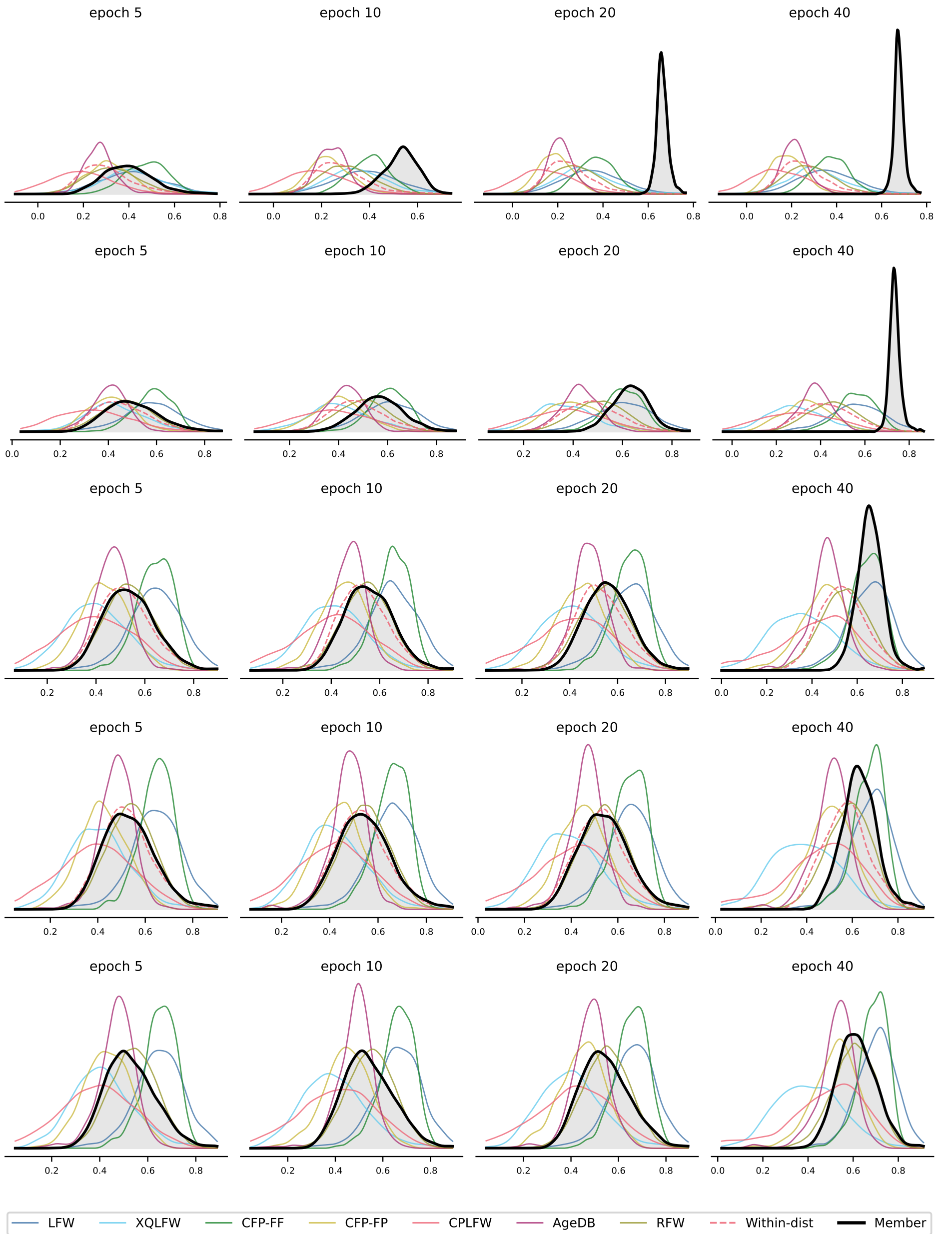


Figure 125. **Vary training duration — IResNet-50 / ArcFace.** Rows: $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

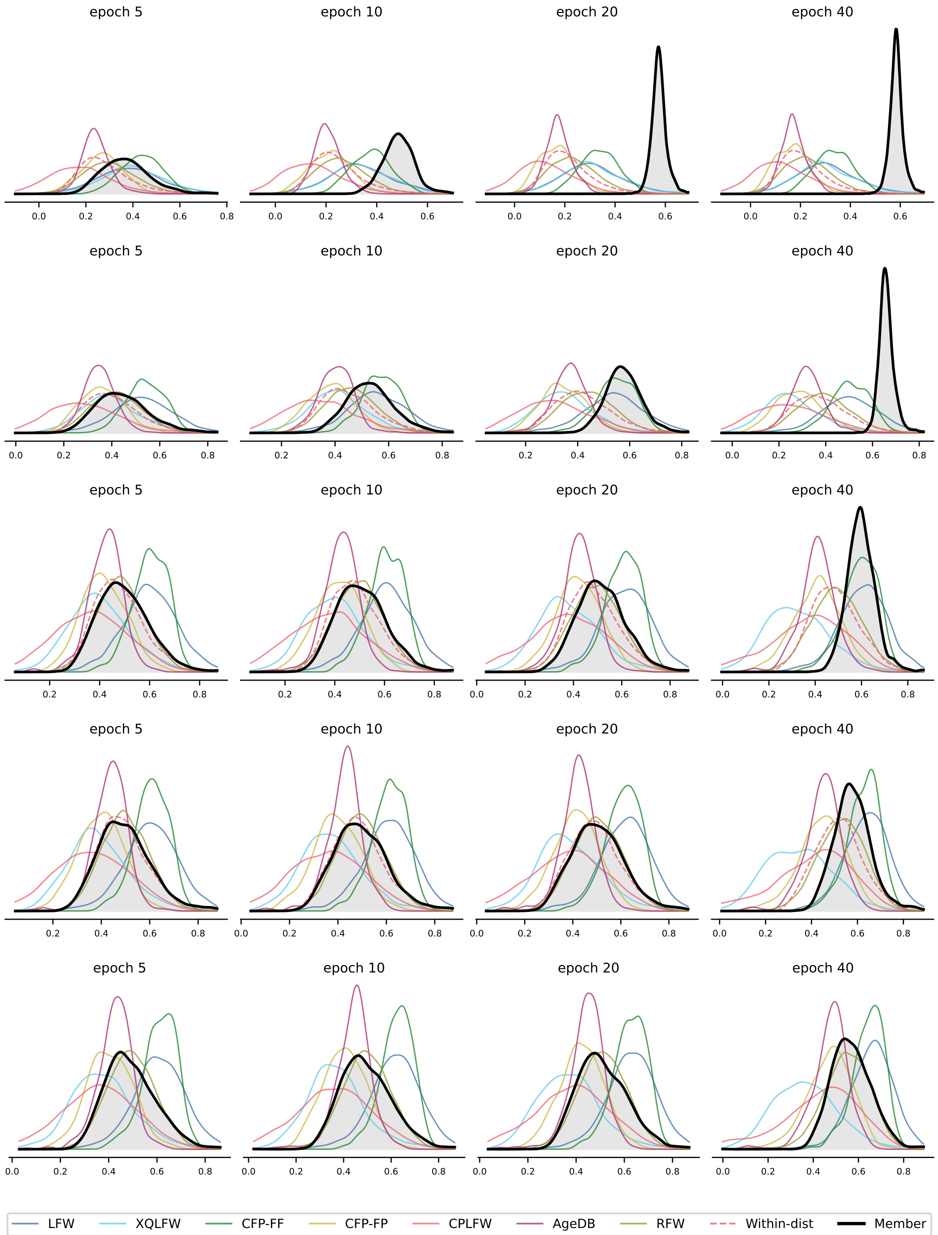


Figure 126. **Vary training duration — IResNet-50 / CosFace.** Rows: $n_{\text{ids}} \in \{1\text{K},\ 10\text{K},\ 50\text{K},\ 100\text{K},\ \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

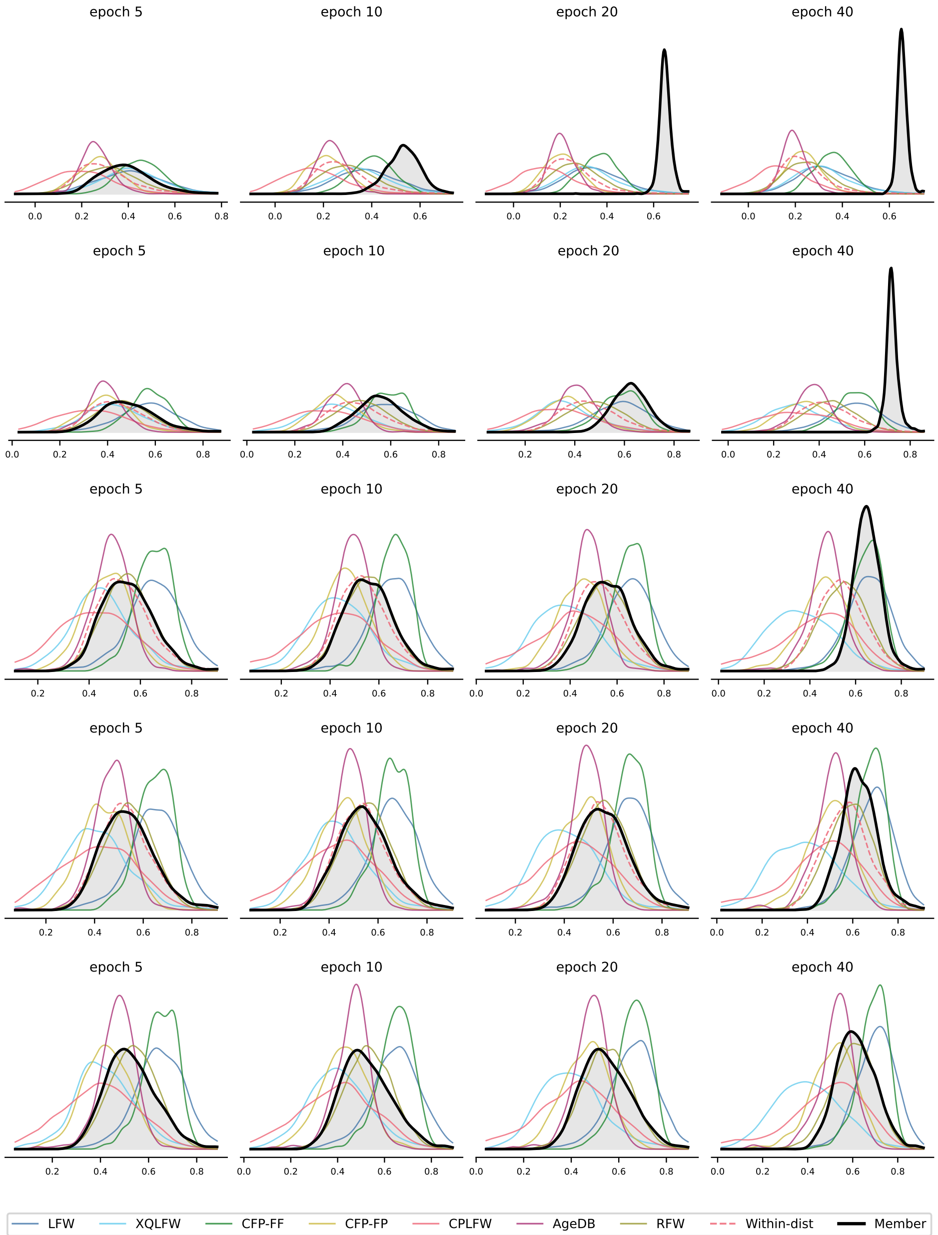


Figure 127. **Vary training duration — IResNet-50 / MagFace.** Rows: $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

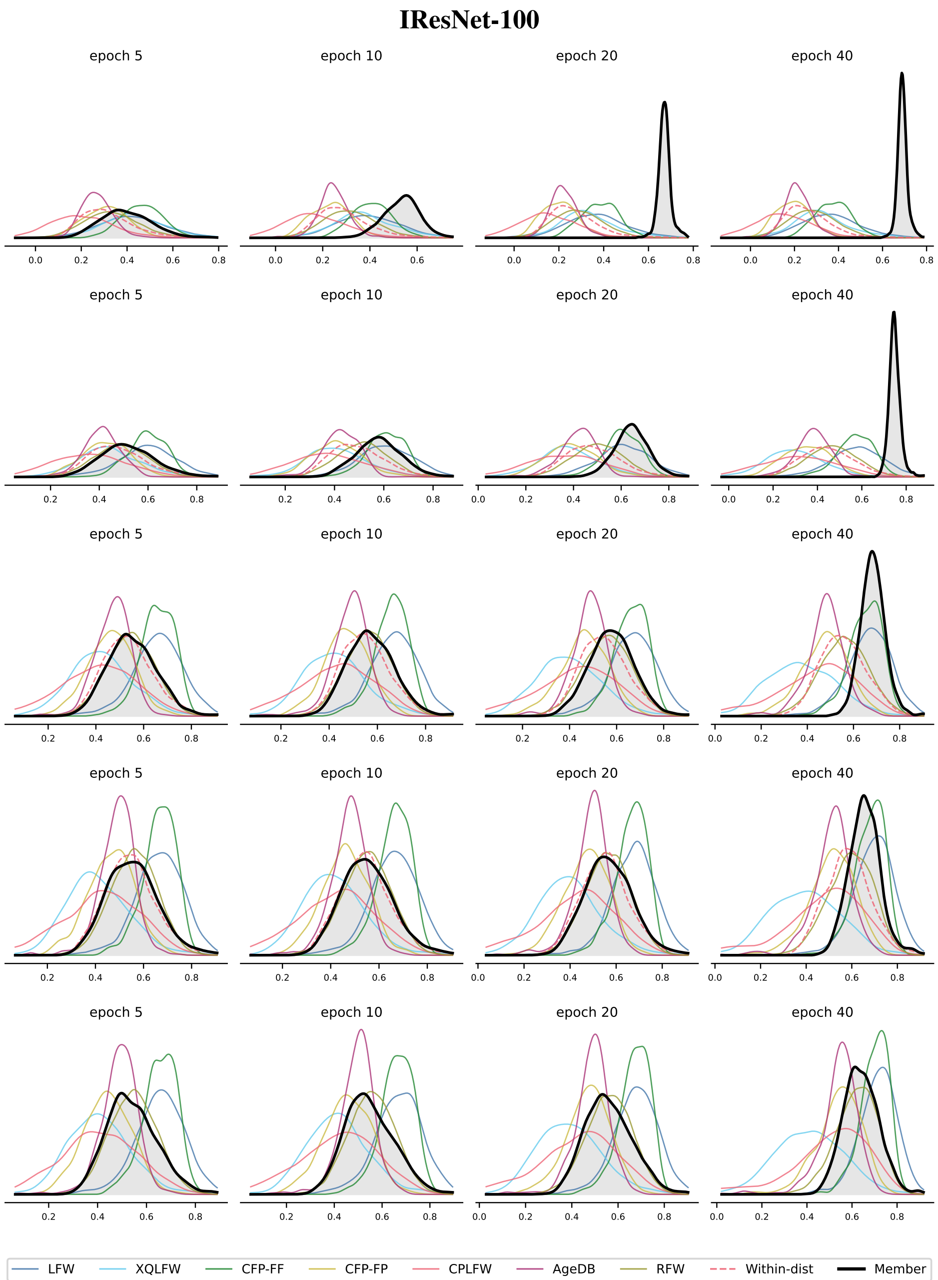


Figure 128. **Vary training duration — IResNet-100 / ArcFace.** Rows: $n_{\text{ids}} \in \{1\text{K},\ 10\text{K},\ 50\text{K},\ 100\text{K},\ \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

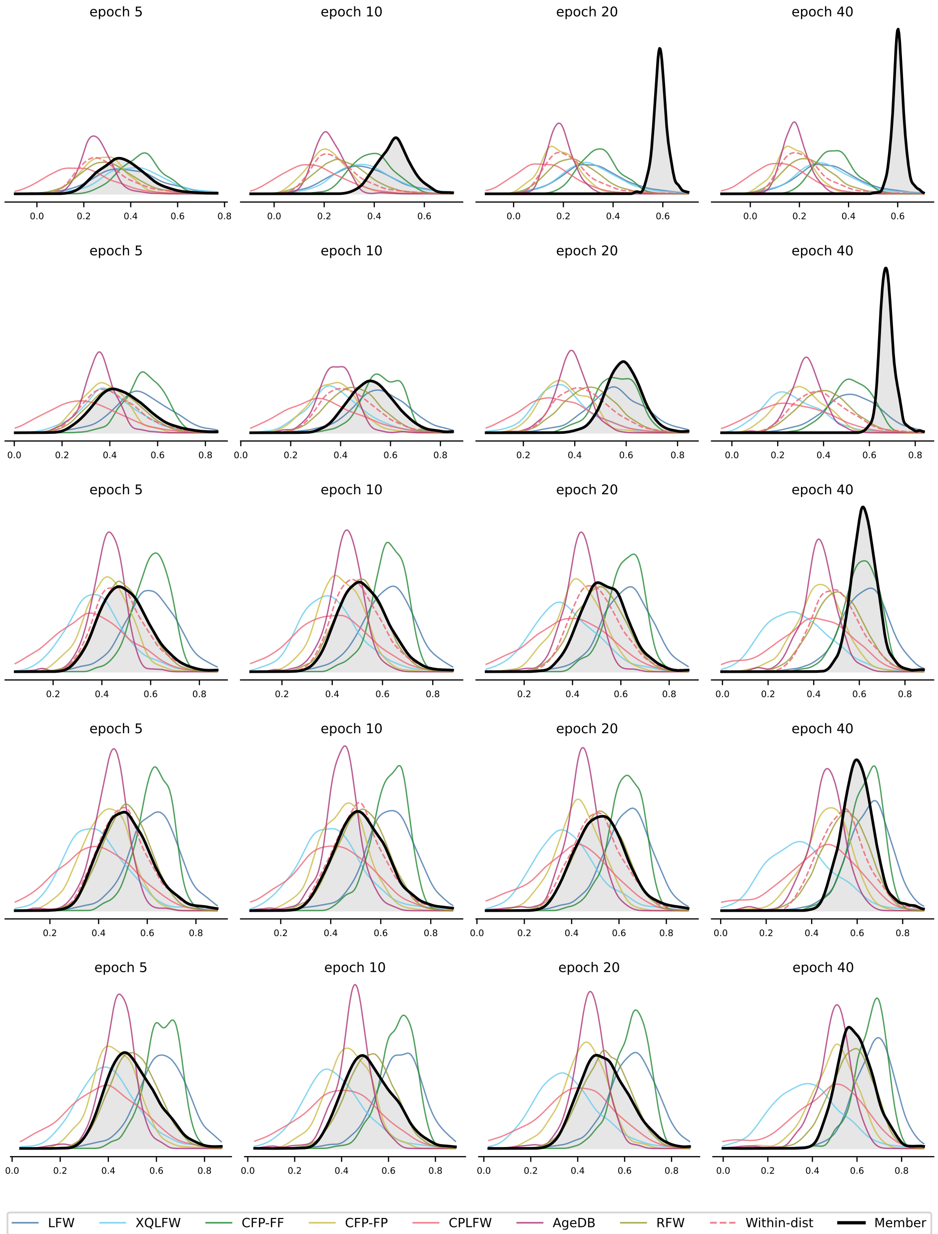


Figure 129. **Vary training duration — IResNet-100 / CosFace.** Rows: $n_{\text{ids}} \in \{1\text{K},\ 10\text{K},\ 50\text{K},\ 100\text{K},\ \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

4.2. Vary training-set size

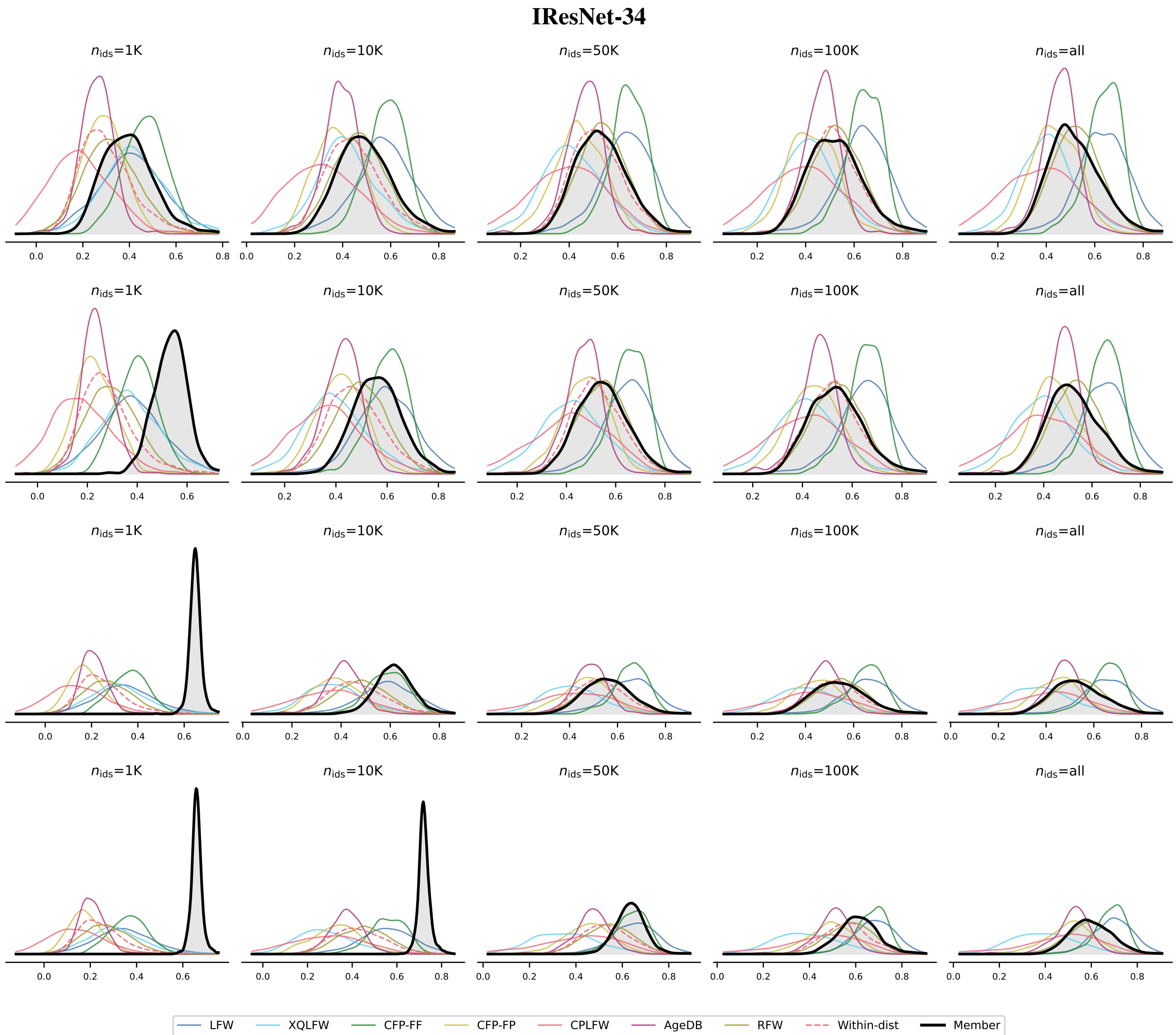


Figure 130. **Vary training-set size — IResNet-34 / ArcFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

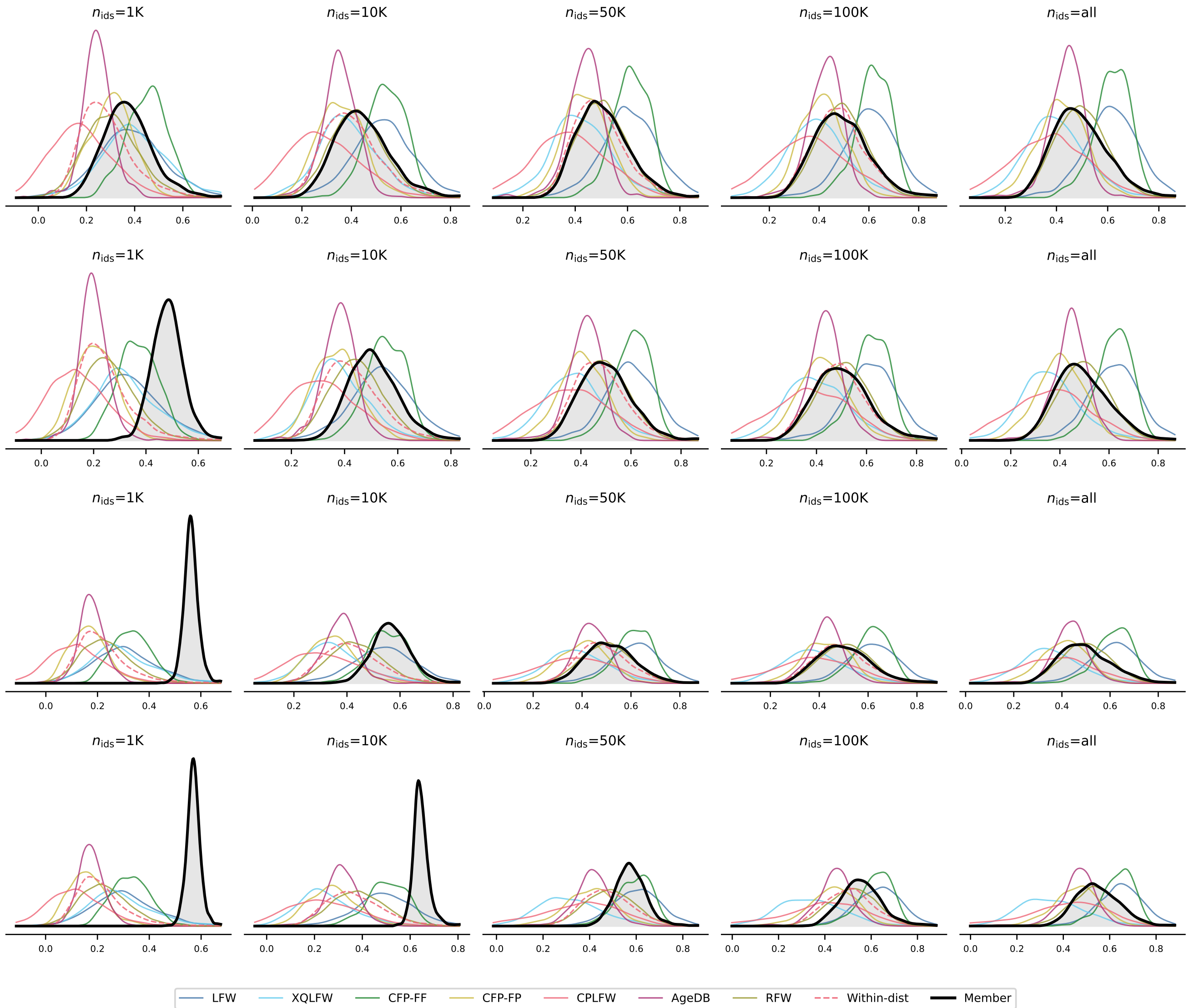


Figure 131. **Vary training-set size — IResNet-34 / CosFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

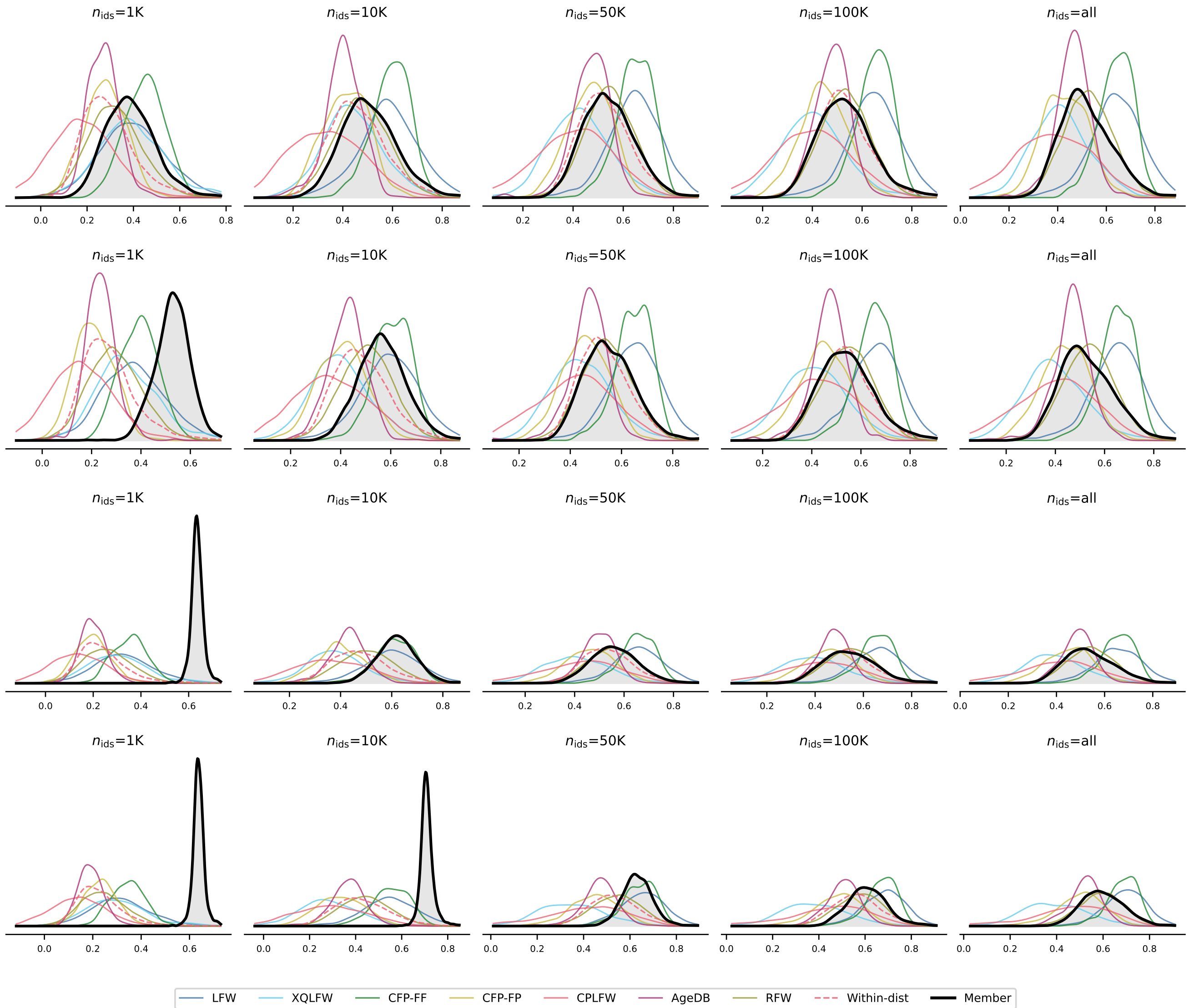


Figure 132. **Vary training-set size — IResNet-34 / MagFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

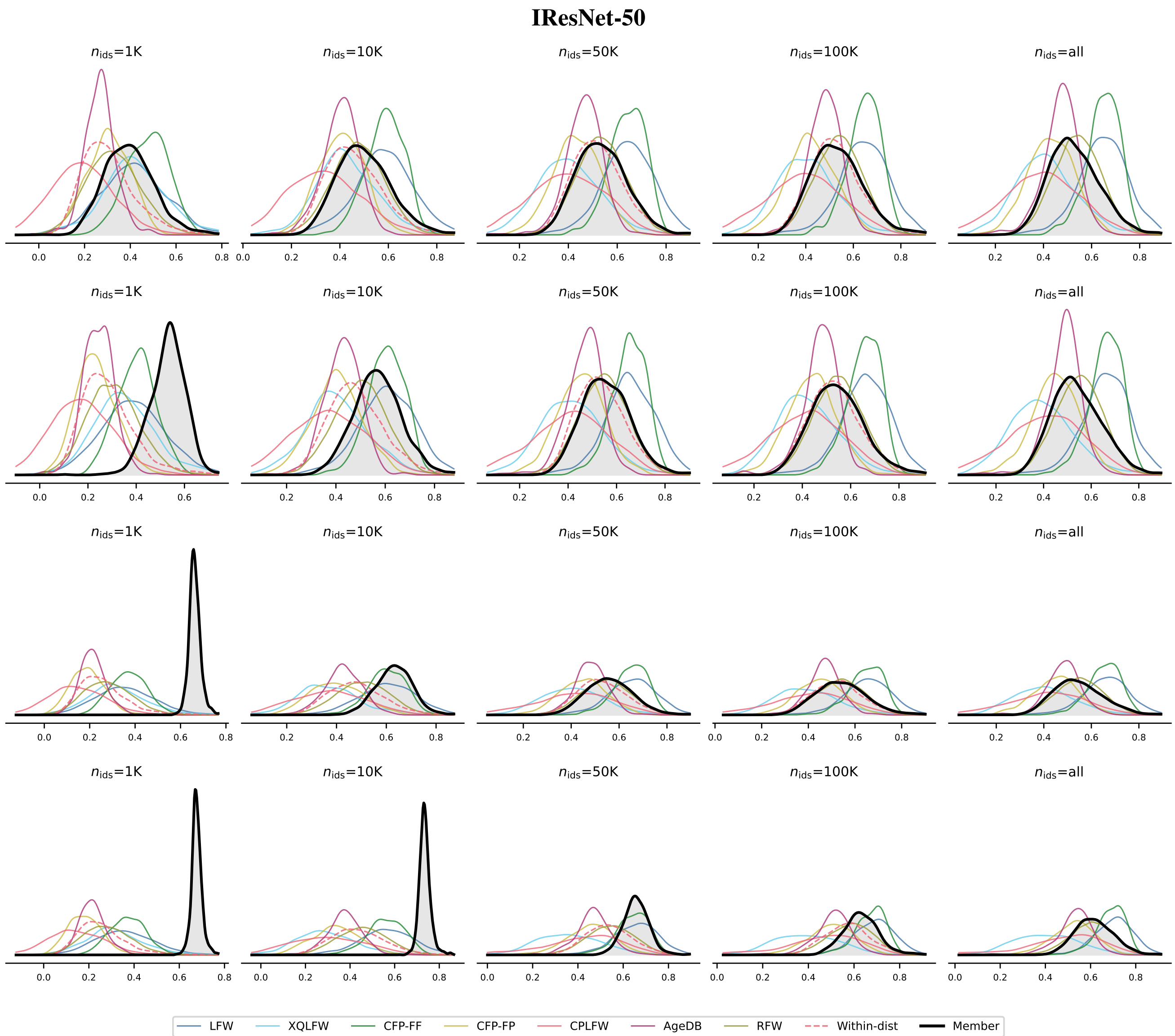


Figure 133. **Vary training-set size — IResNet-50 / ArcFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

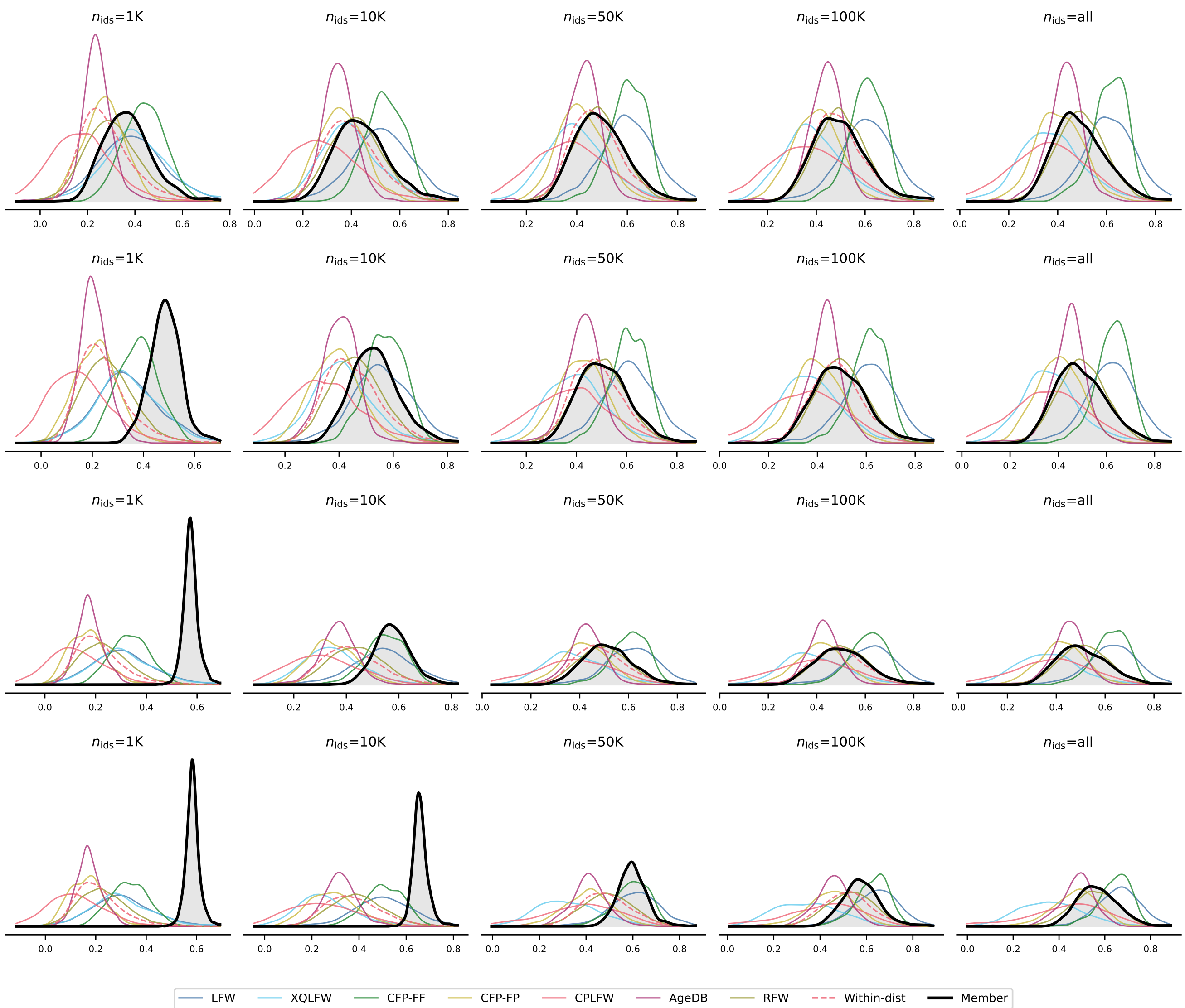


Figure 134. **Vary training-set size — IResNet-50 / CosFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

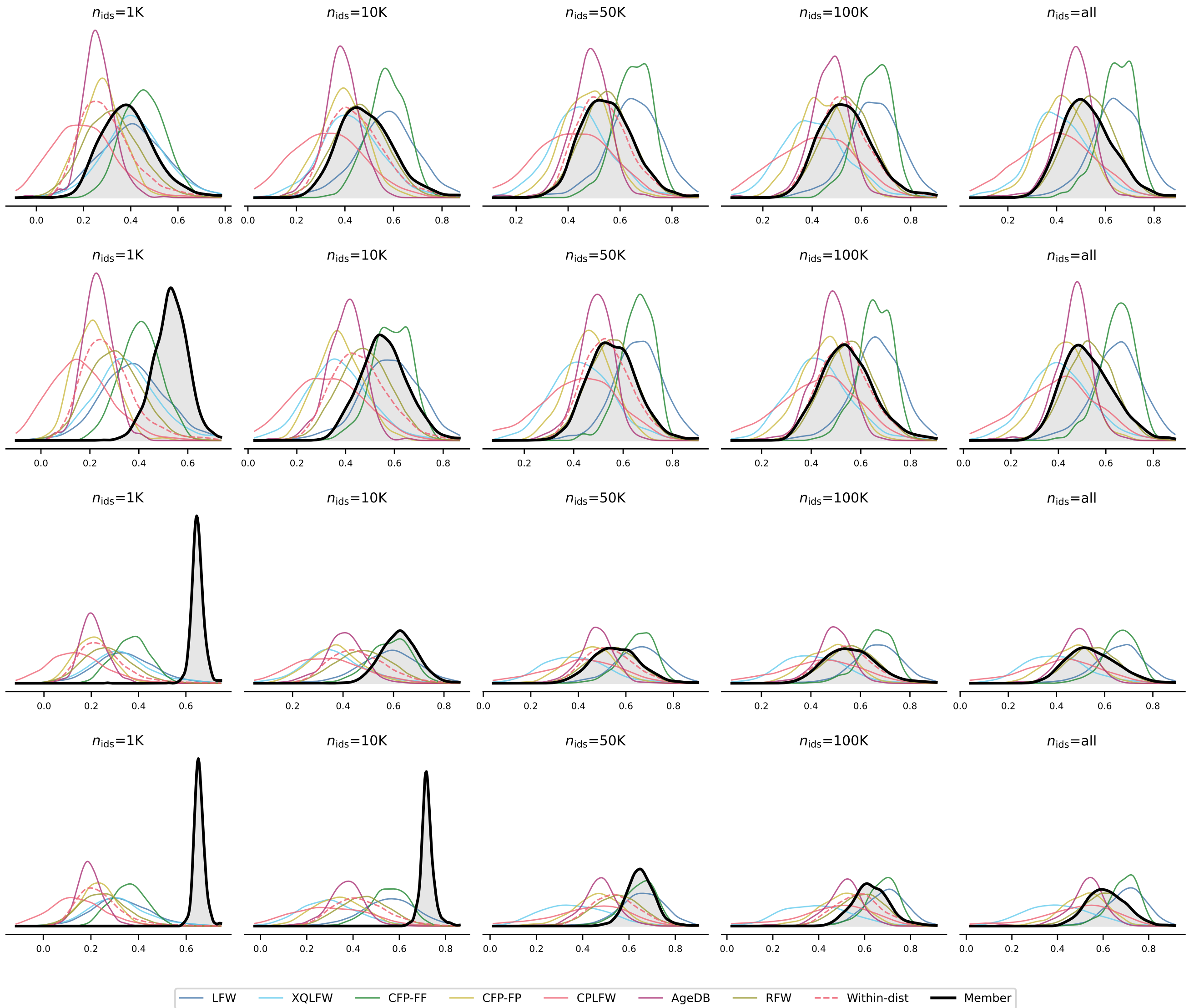


Figure 135. **Vary training-set size — IResNet-50 / MagFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

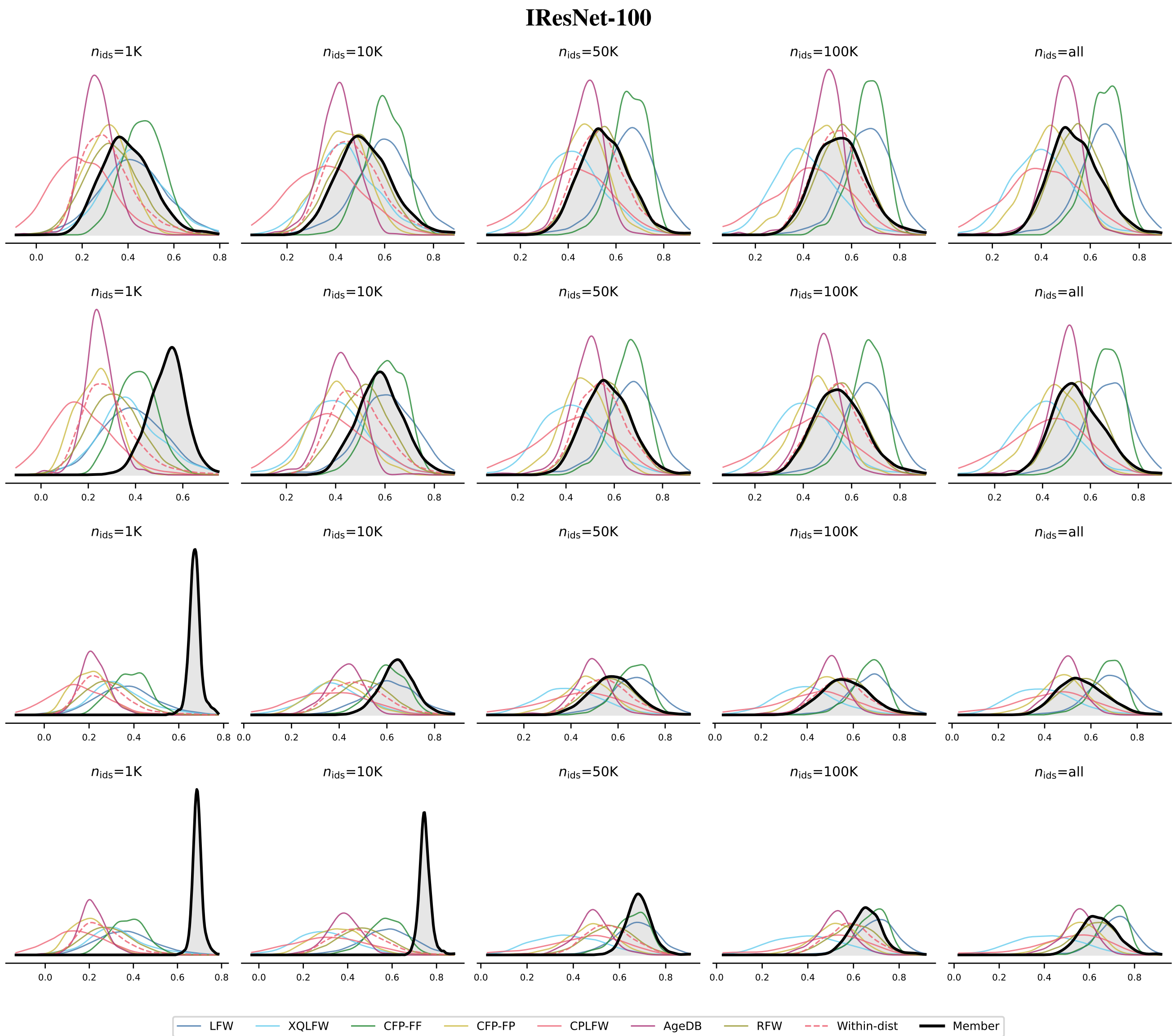


Figure 136. **Vary training-set size — IResNet-100 / ArcFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

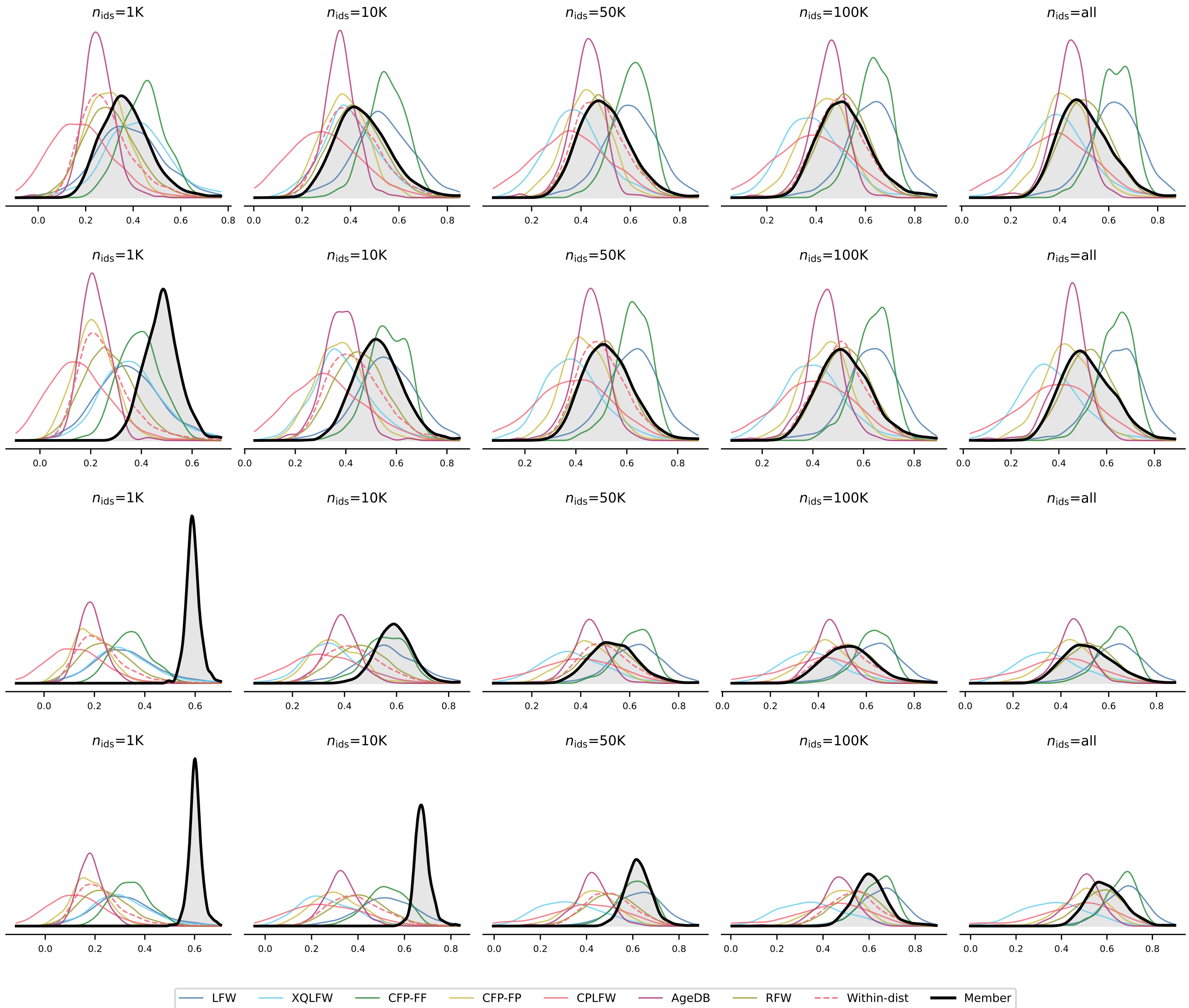


Figure 137. **Vary training-set size — IResNet-100 / CosFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

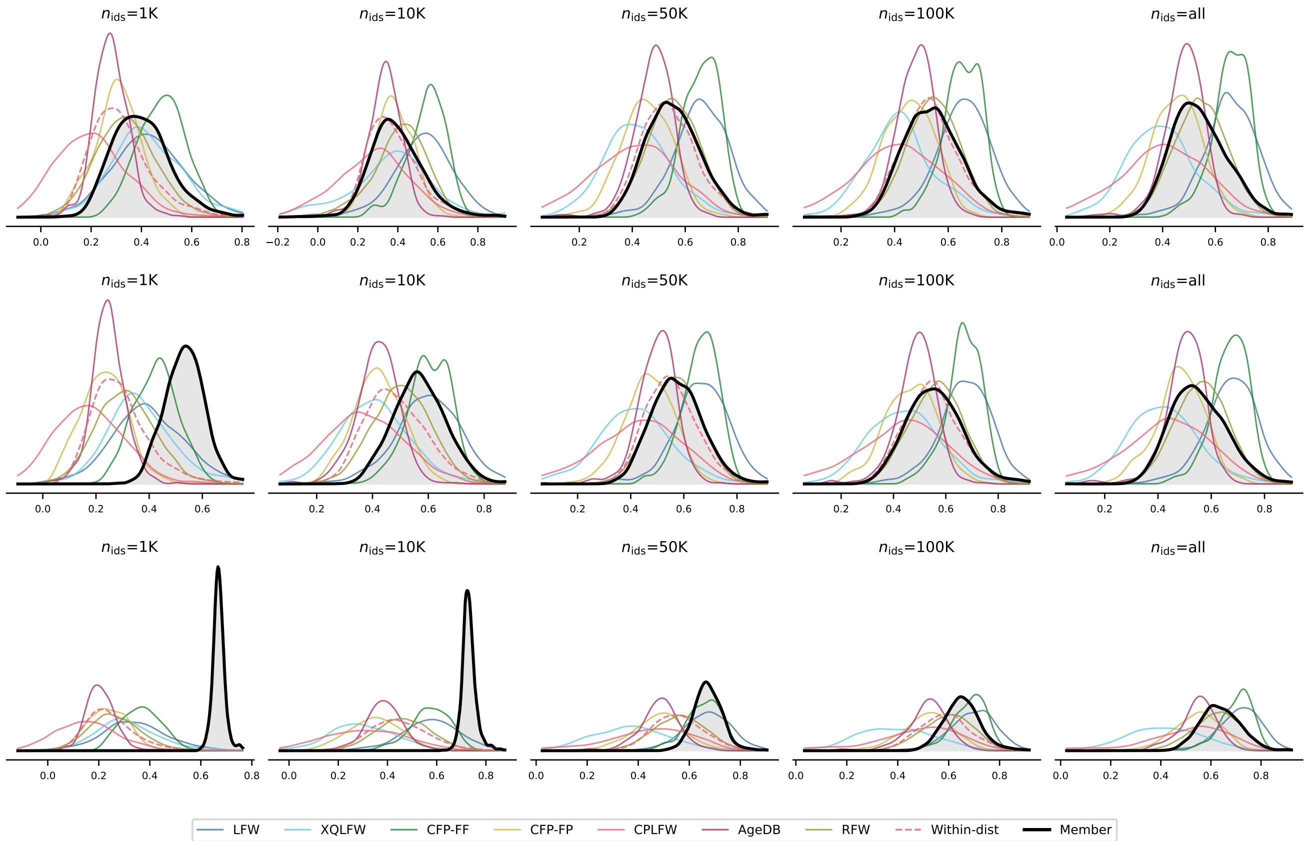


Figure 138. **Vary training-set size — IResNet-100 / MagFace.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

4.3. Vary loss head

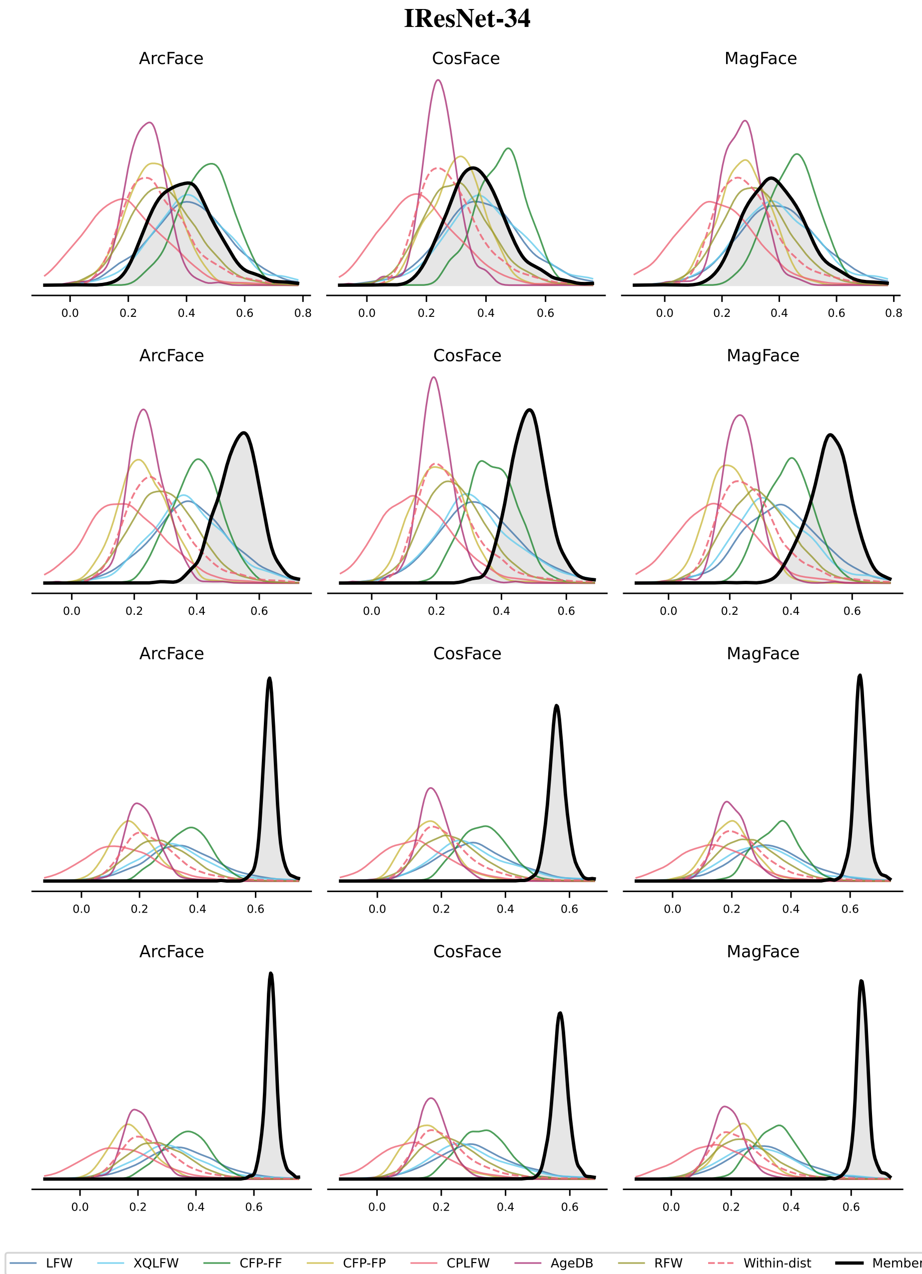


Figure 139. **Vary loss head — IResNet-34 /** $n_{\text{ids}} = 1K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

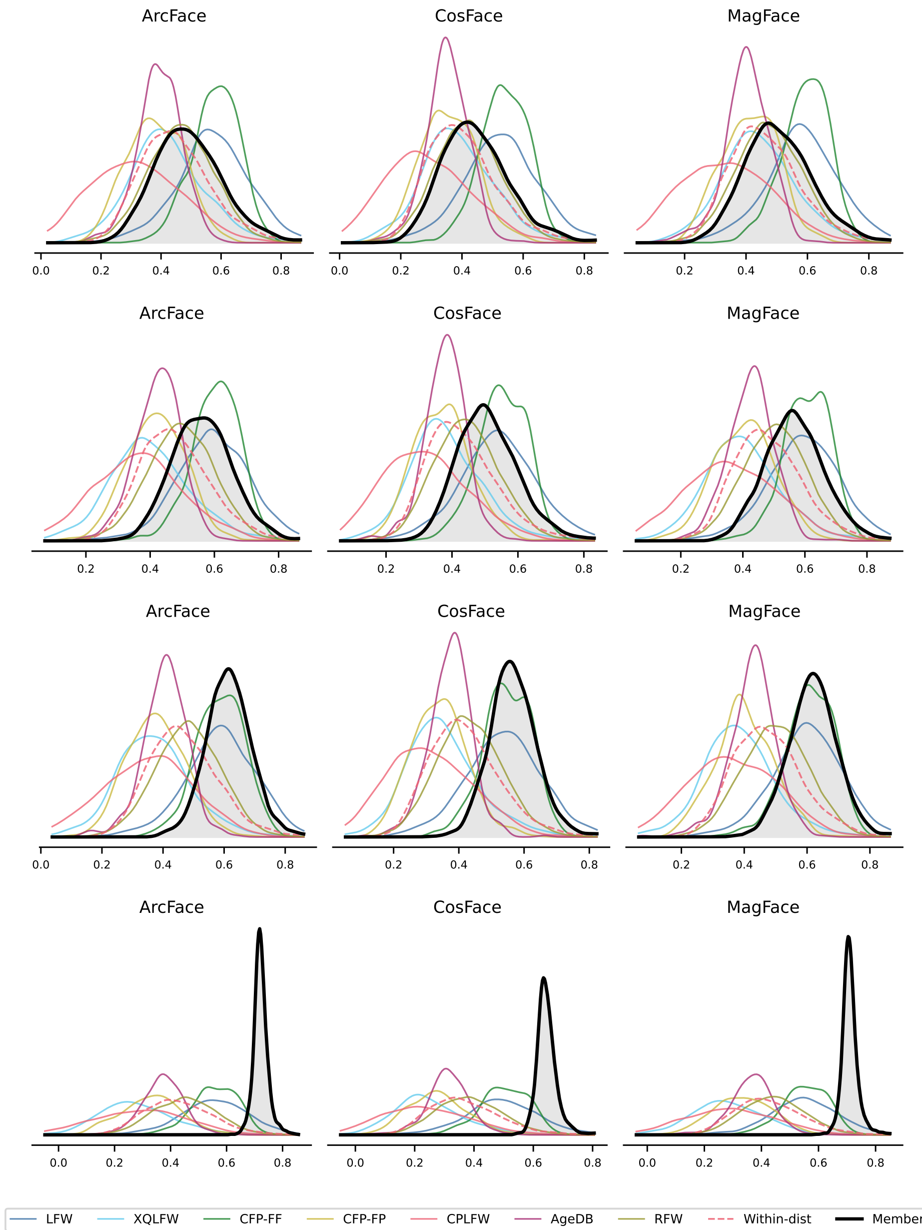


Figure 140. **Vary loss head — IResNet-34 / $n_{\text{ids}} = 10K$.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

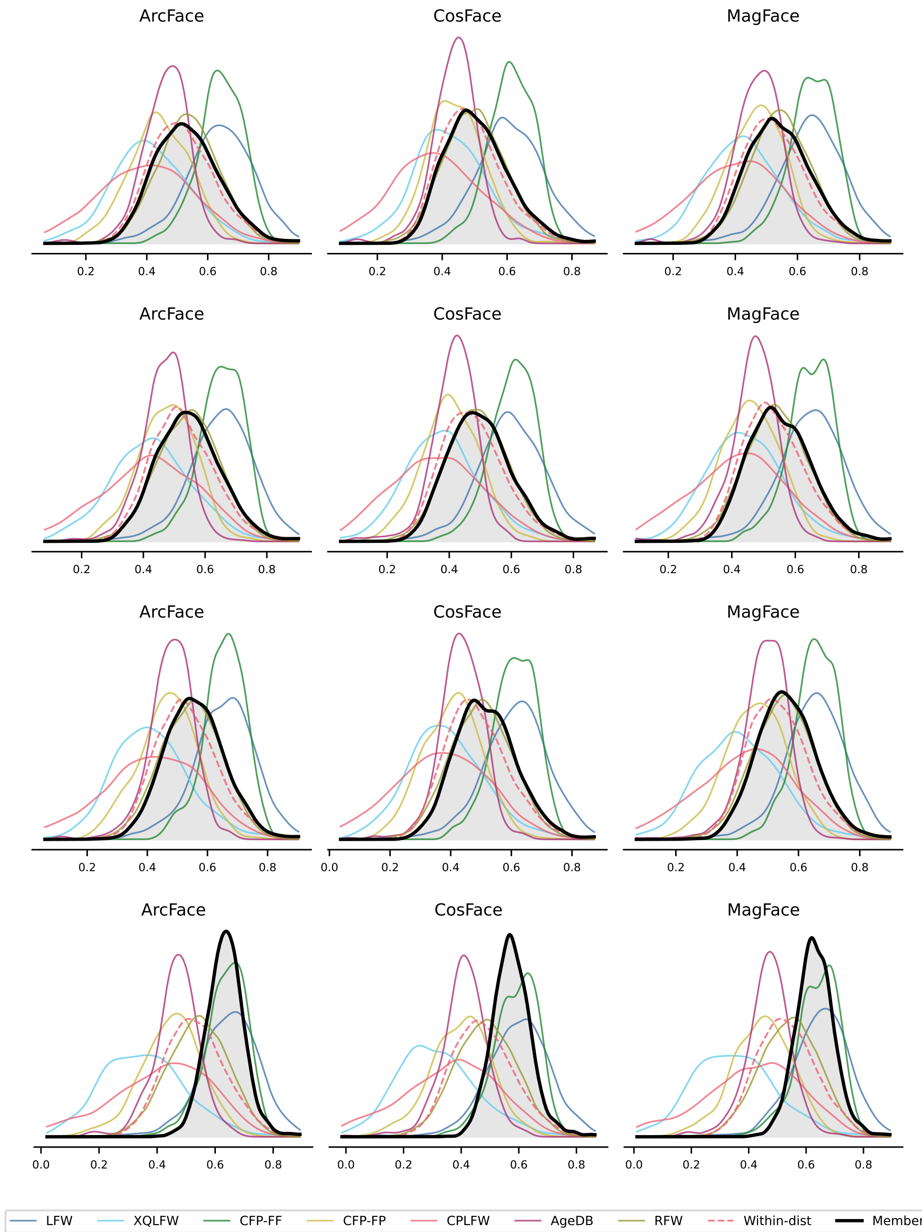


Figure 141. **Vary loss head — IResNet-34 /** $n_{\text{ids}} = 50K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

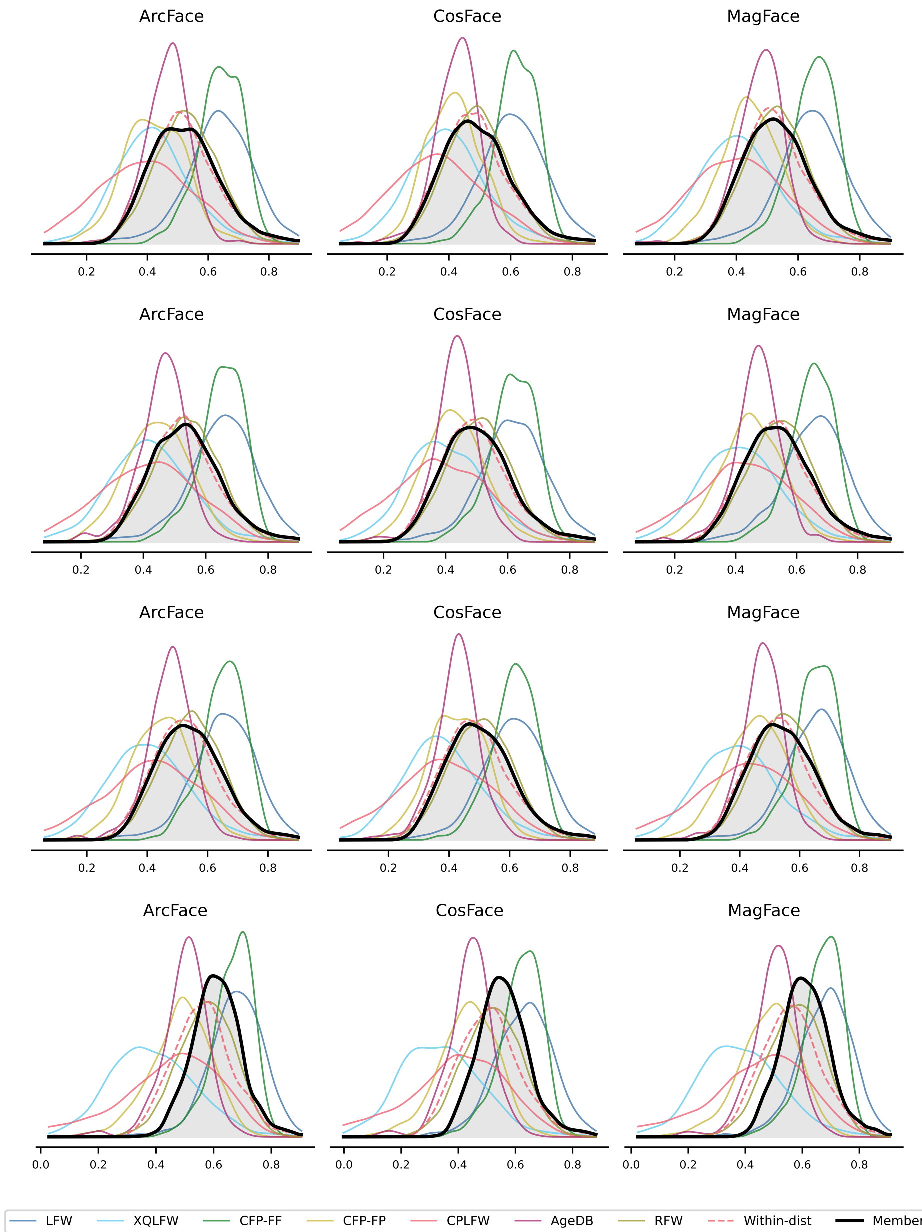


Figure 142. **Vary loss head — IResNet-34 /** $n_{\text{ids}} = 100K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

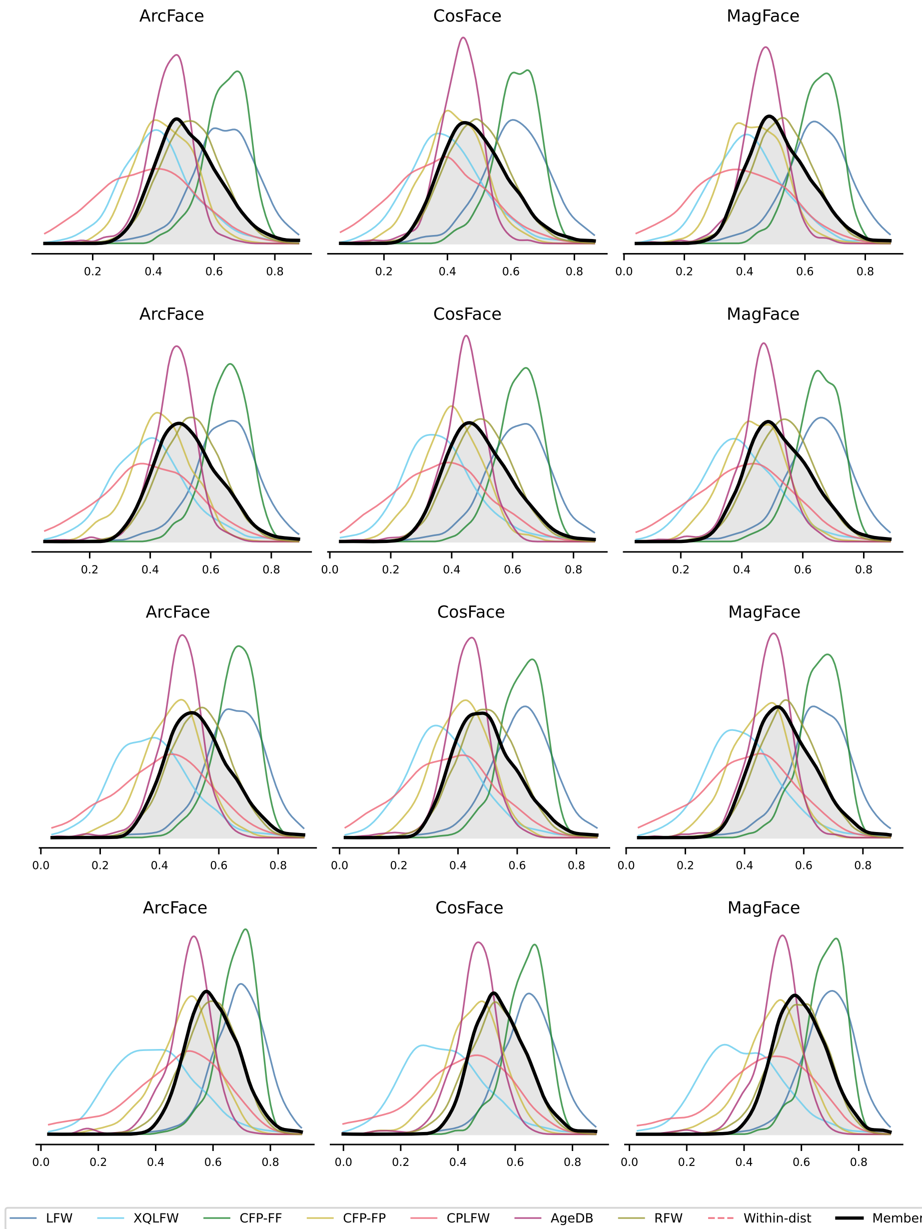


Figure 143. **Vary loss head — IResNet-34 /** $n_{\textbf{ids}} = All$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

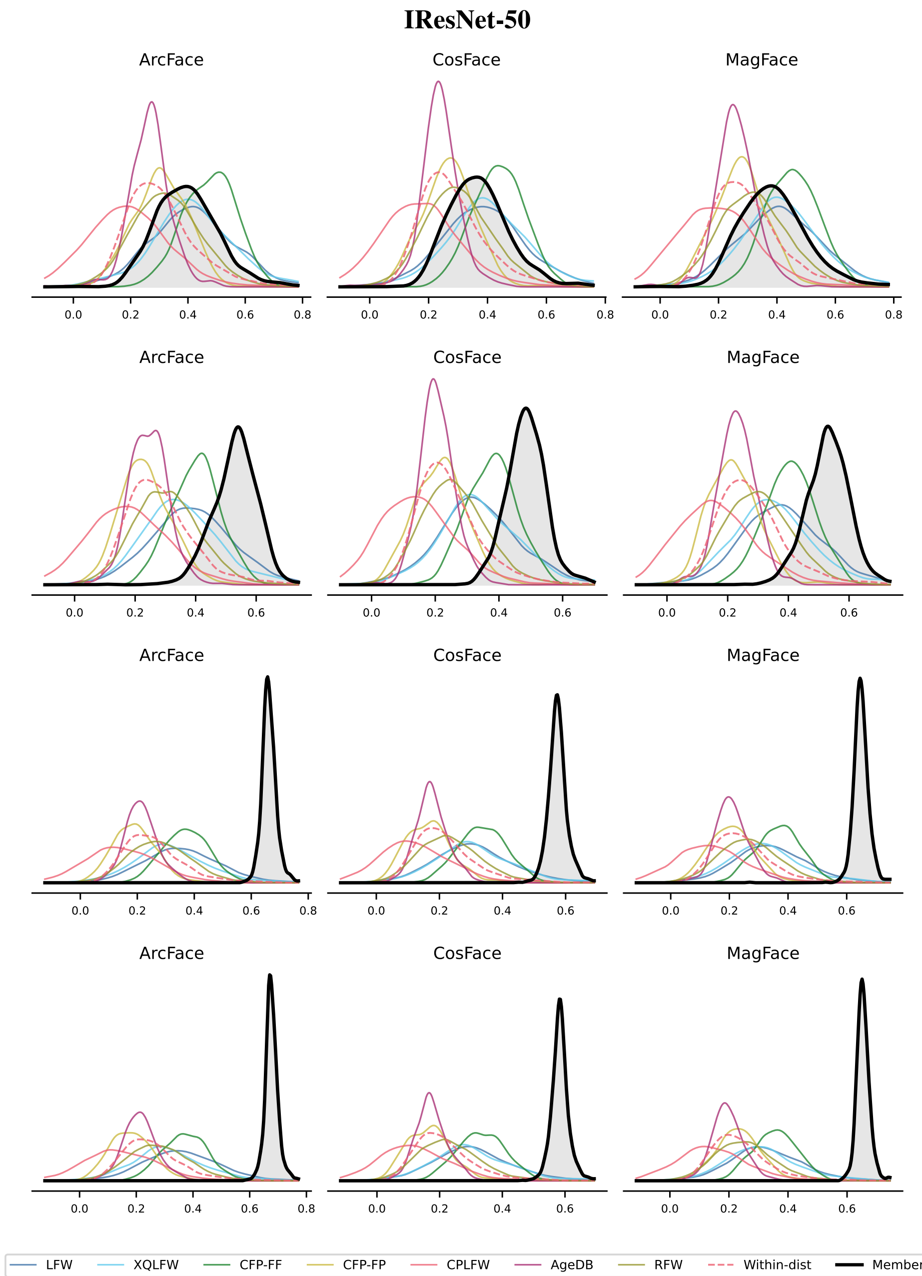


Figure 144. **Vary loss head — IResNet-50 /** $n_{\textbf{ids}} = 1K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

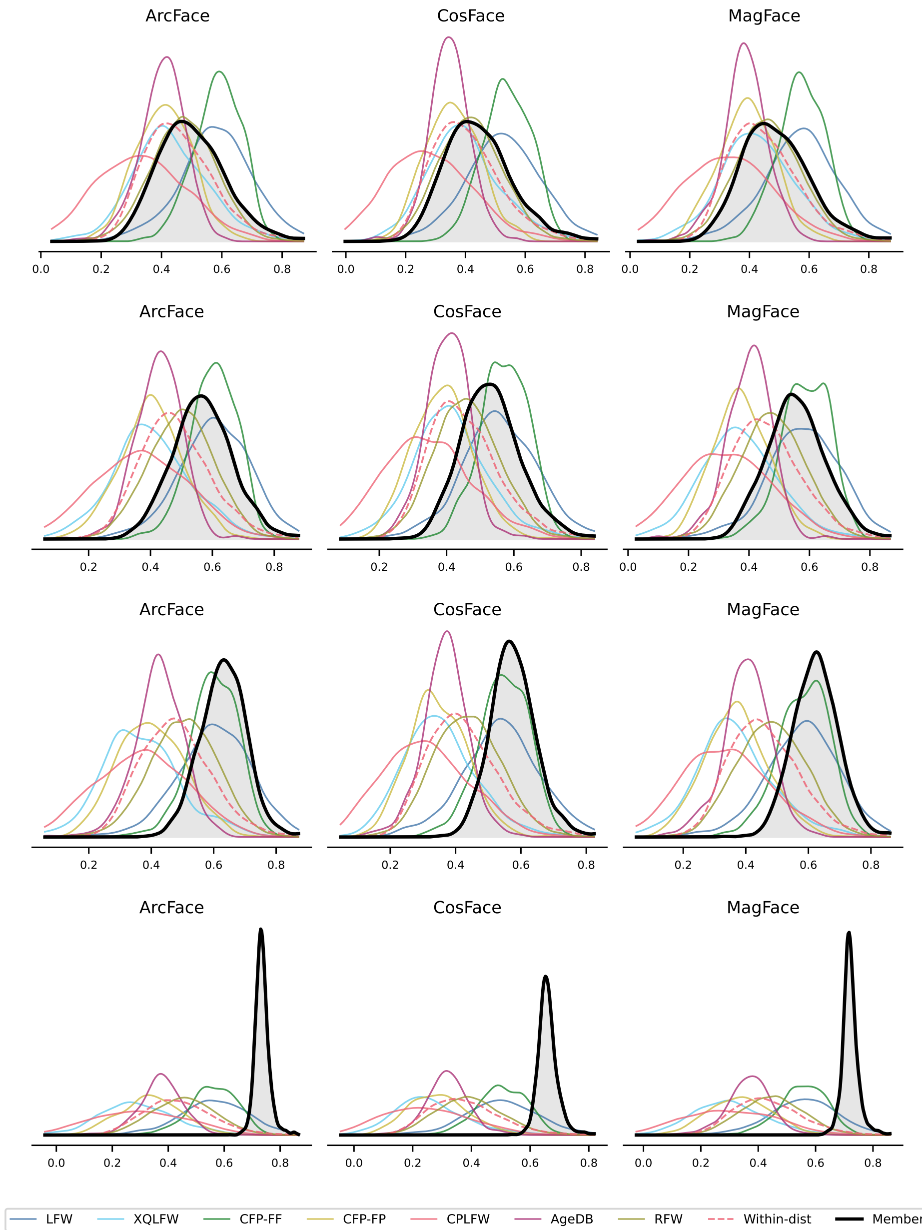


Figure 145. **Vary loss head — IResNet-50 /** $n_{\text{ids}} = 10K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

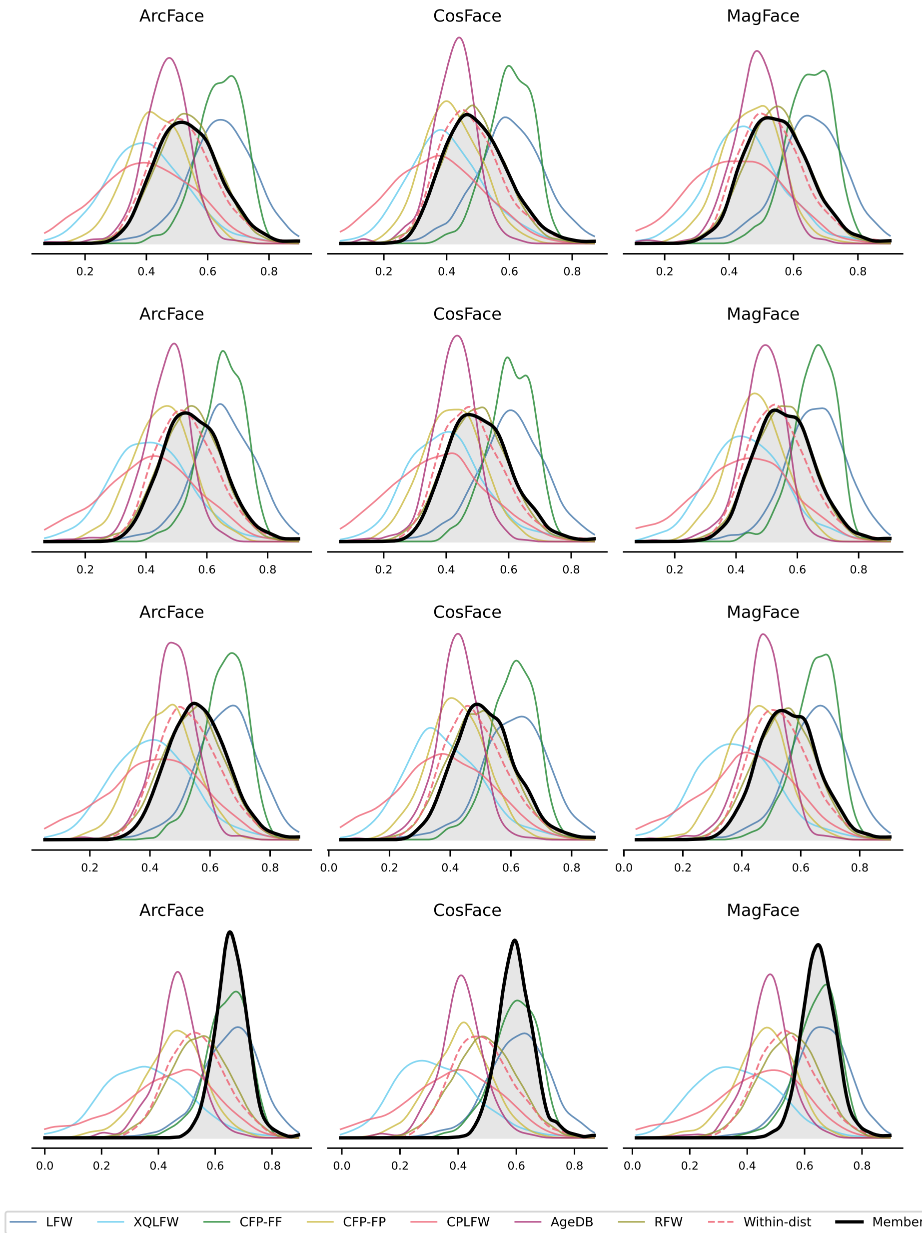


Figure 146. **Vary loss head — IResNet-50 /** $n_{\text{ids}} = 50K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

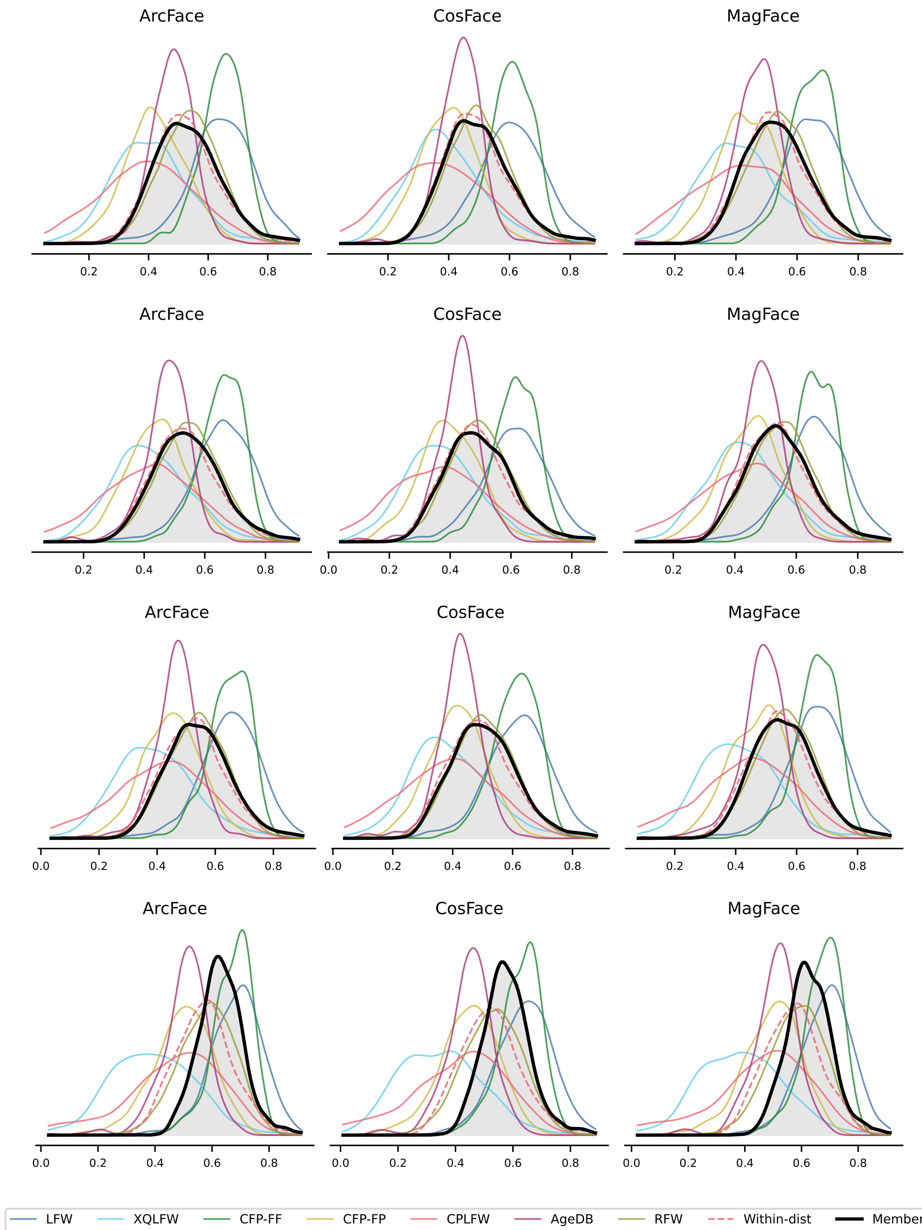


Figure 147. **Vary loss head — IResNet-50 /** $n_{\text{ids}} = 100K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

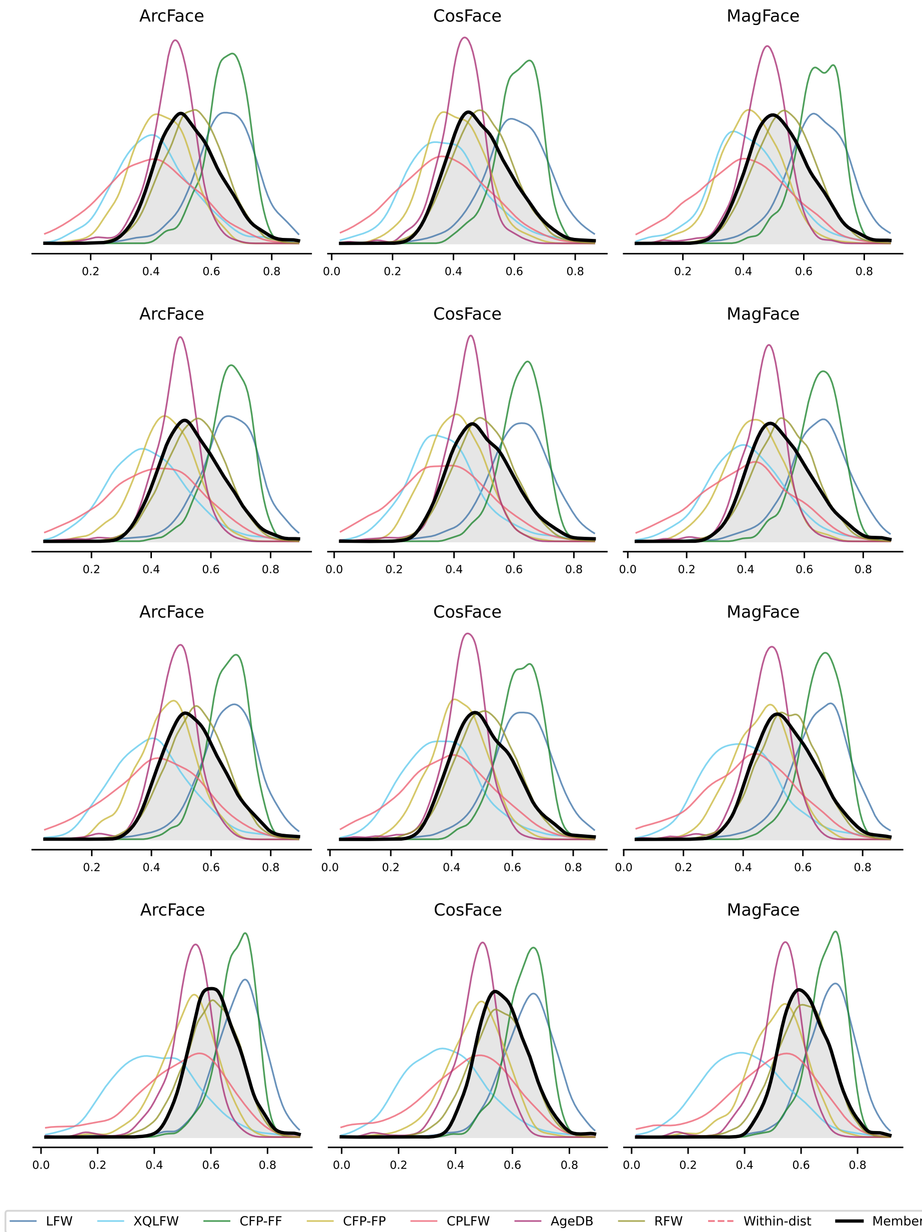


Figure 148. **Vary loss head — IResNet-50 / $n_{\text{ids}} = All$.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

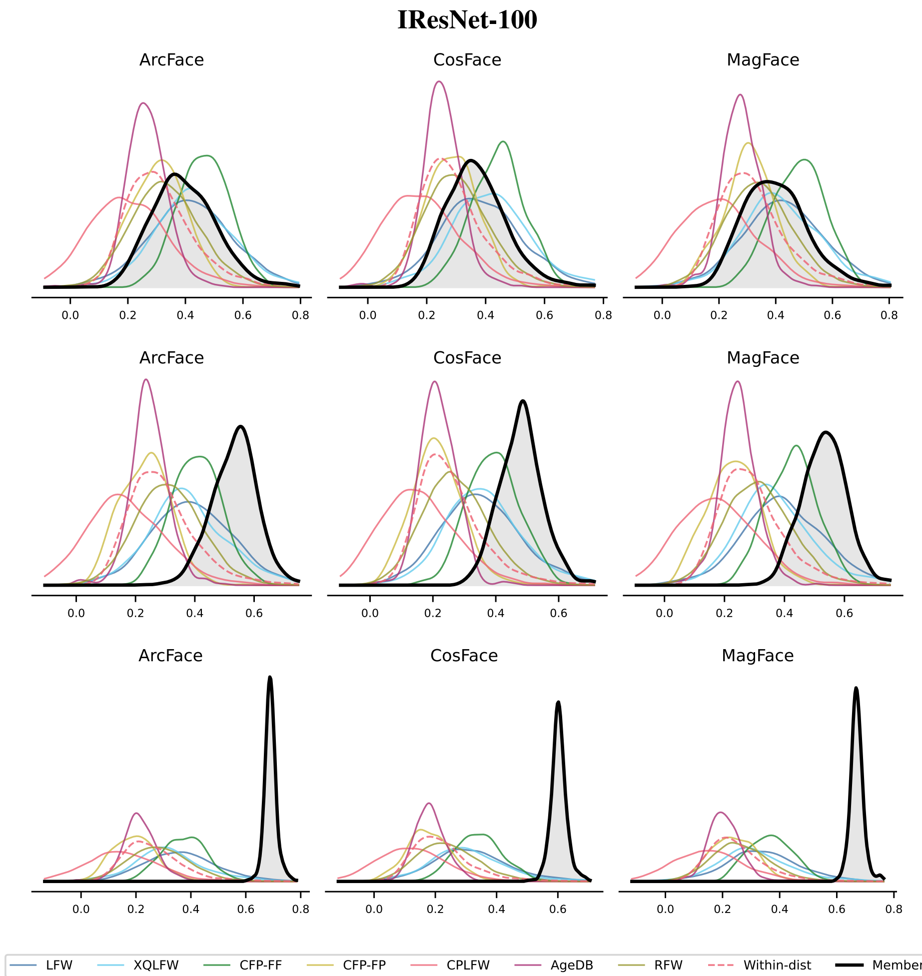


Figure 149. **Vary loss head — IResNet-100 /** $n_{\text{ids}} = 1K$**.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

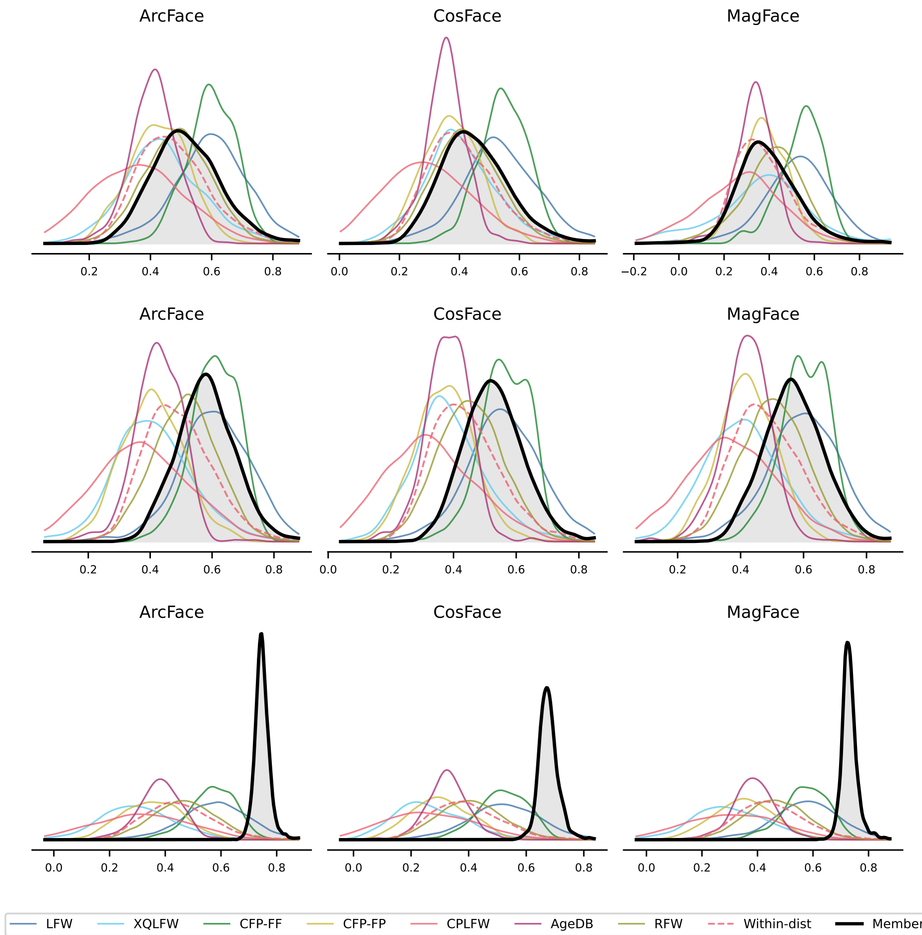


Figure 150. **Vary loss head — IResNet-100 /** $n_{\mathbf{ids}} = 10K$**.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

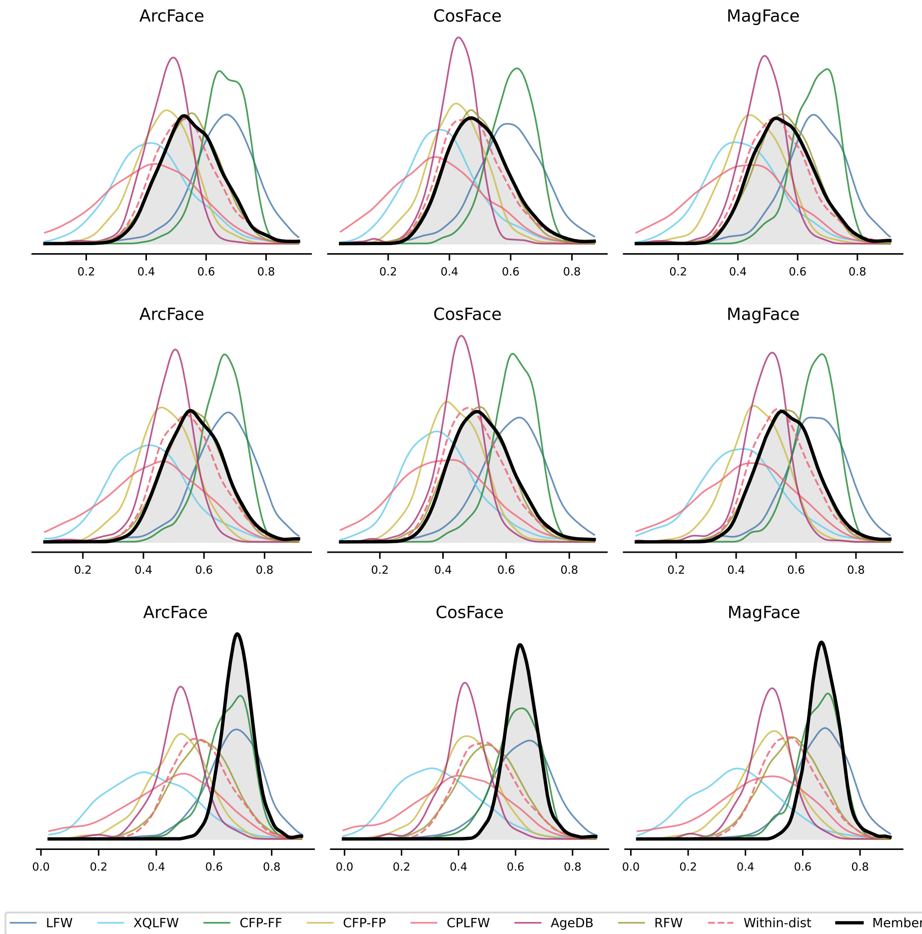


Figure 151. **Vary loss head — IResNet-100 / $n_{\text{ids}} = 50K$.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

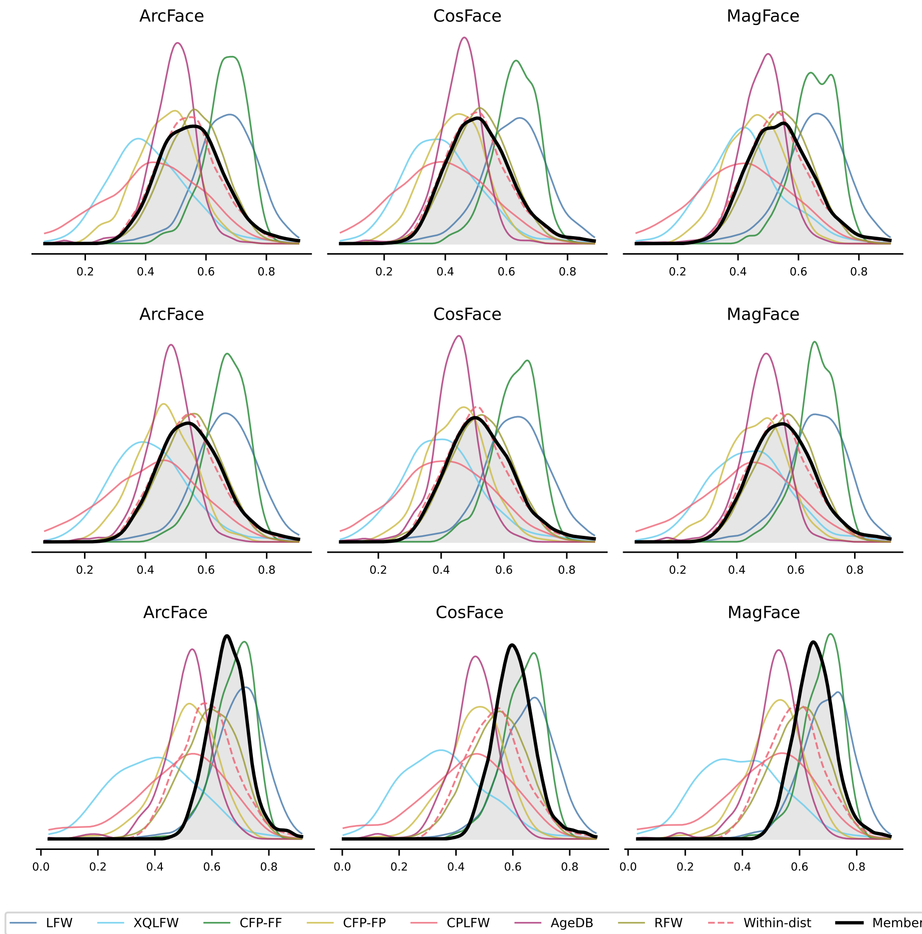


Figure 152. **Vary loss head — IResNet-100 /** $n_{\text{ids}} = 100K$**.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

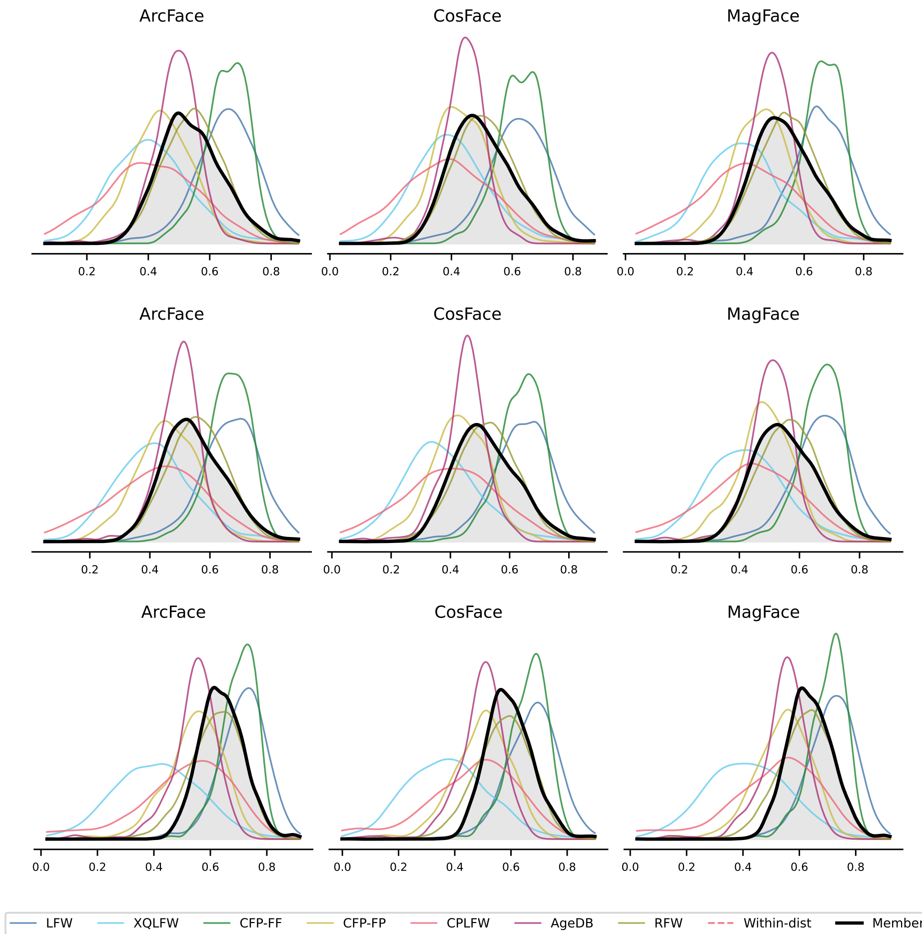


Figure 153. **Vary loss head — IResNet-100 / $n_{\text{ids}} = All$.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

4.4. Vary backbone

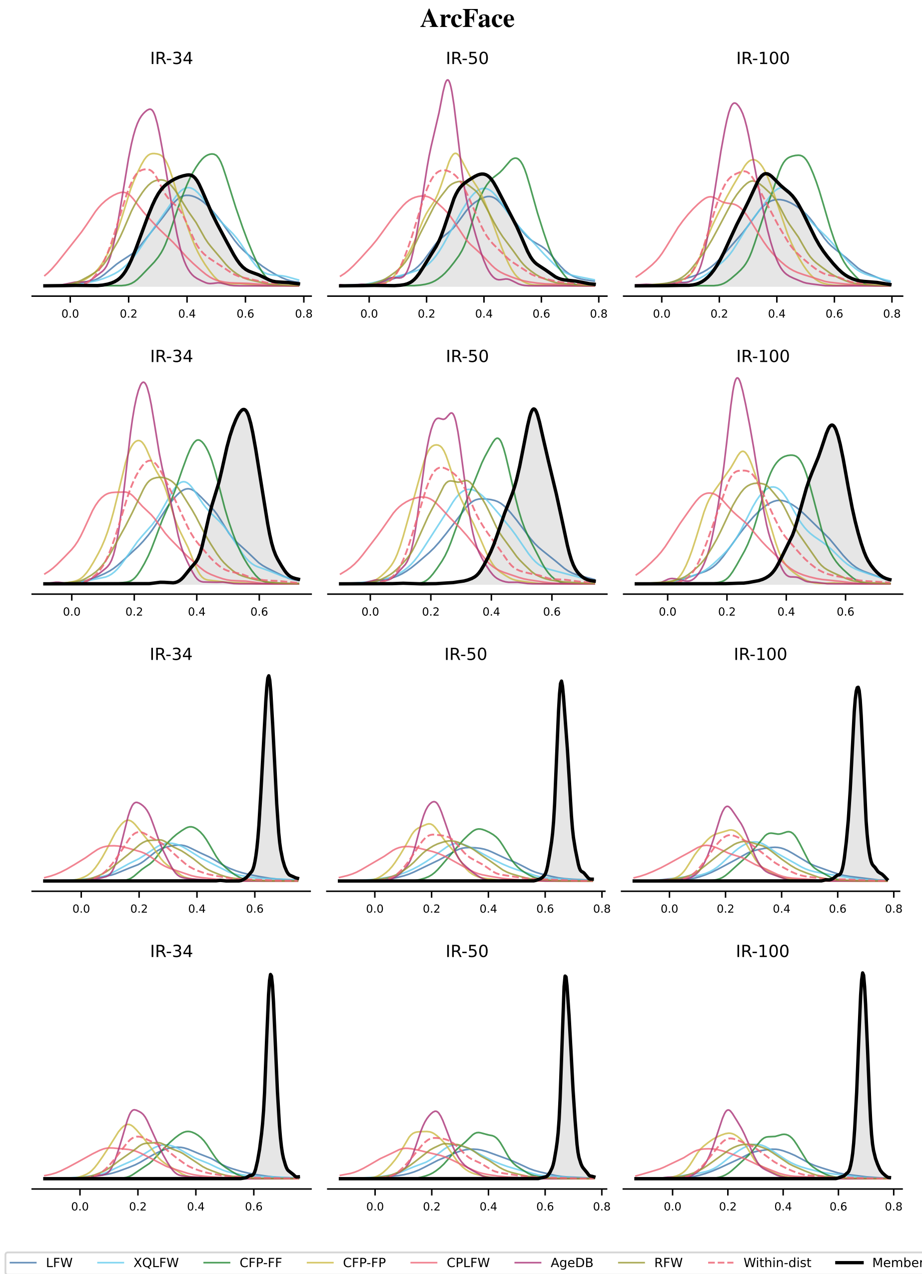


Figure 154. **Vary backbone — ArcFace /** $n_{\text{ids}} = 1K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

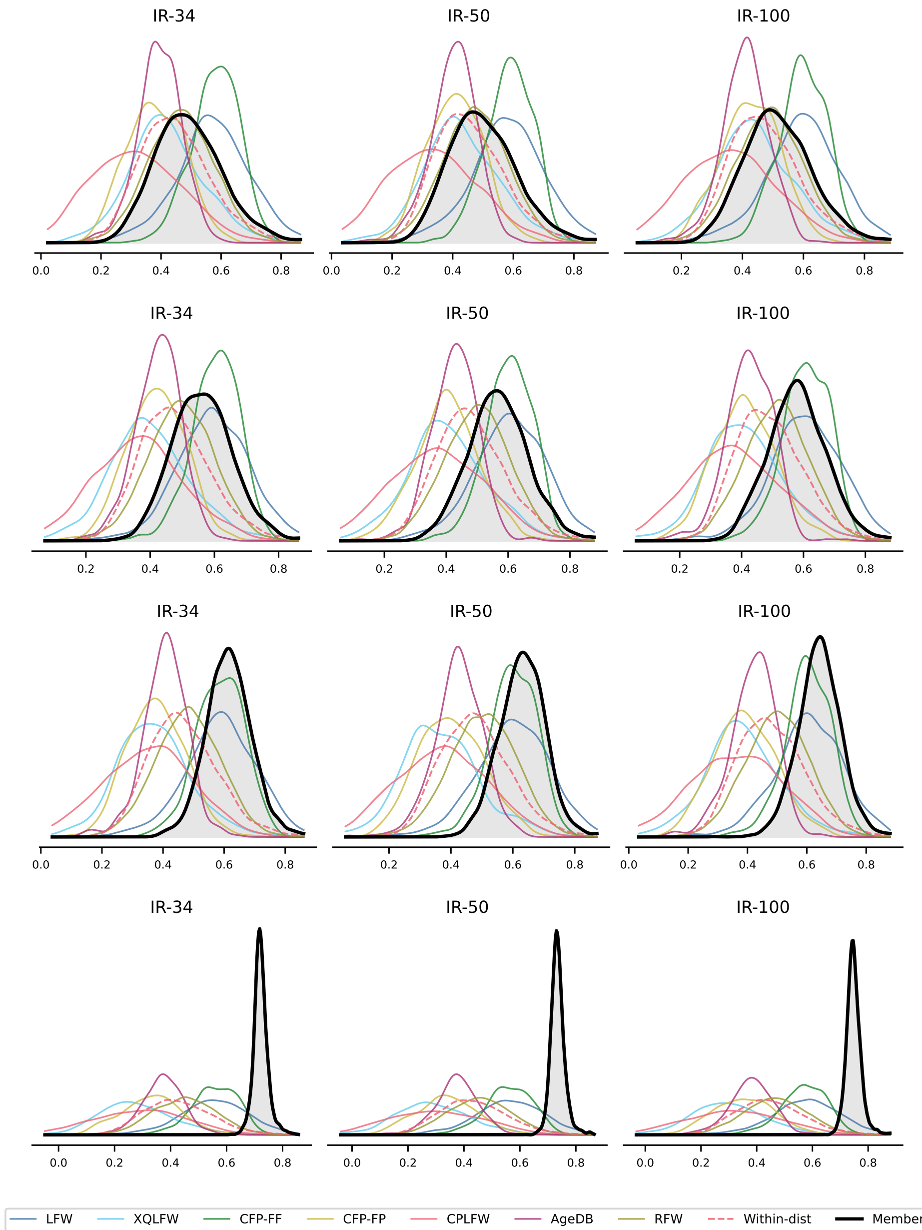


Figure 155. **Vary backbone — ArcFace /** $n_{\textbf{ids}} = 10K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

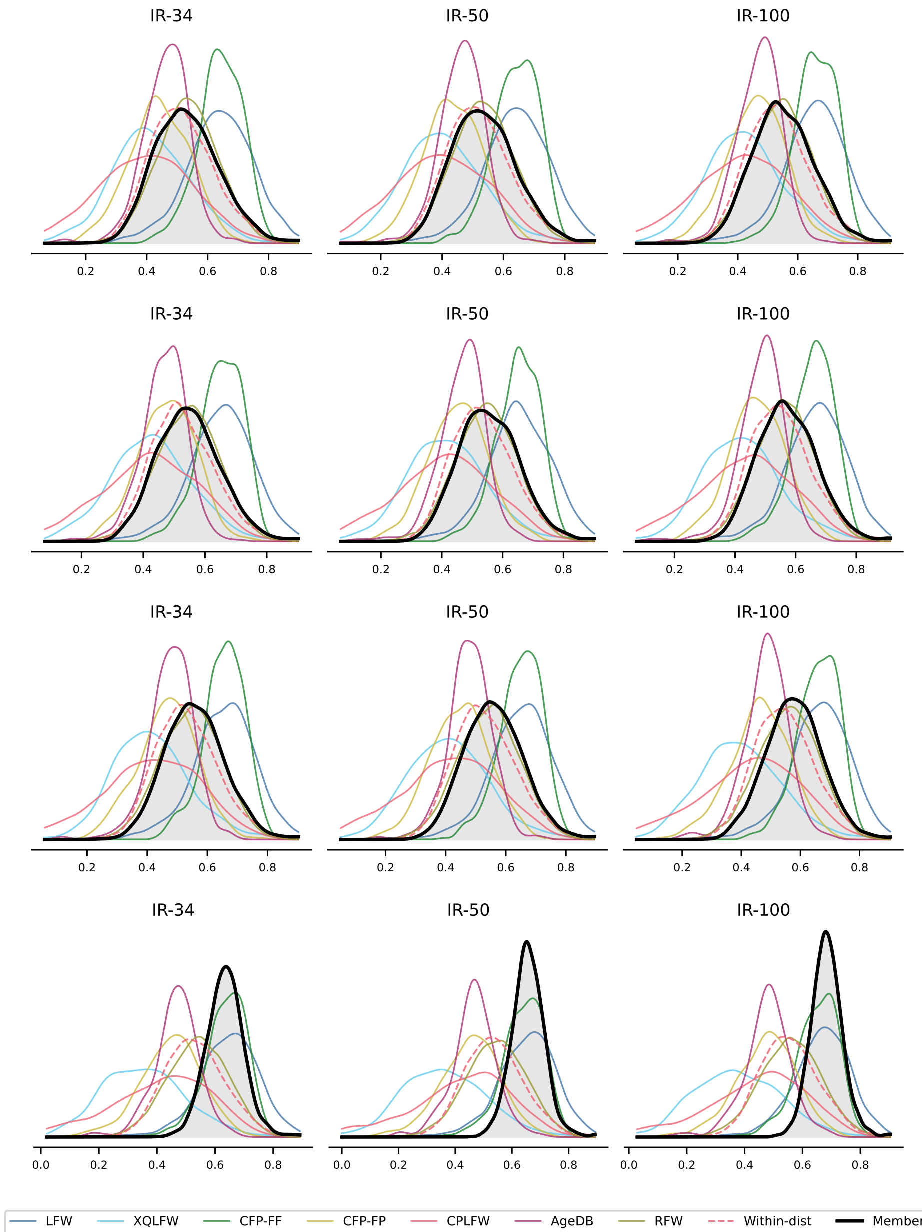


Figure 156. **Vary backbone — ArcFace /** $n_{\text{ids}} = 50K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

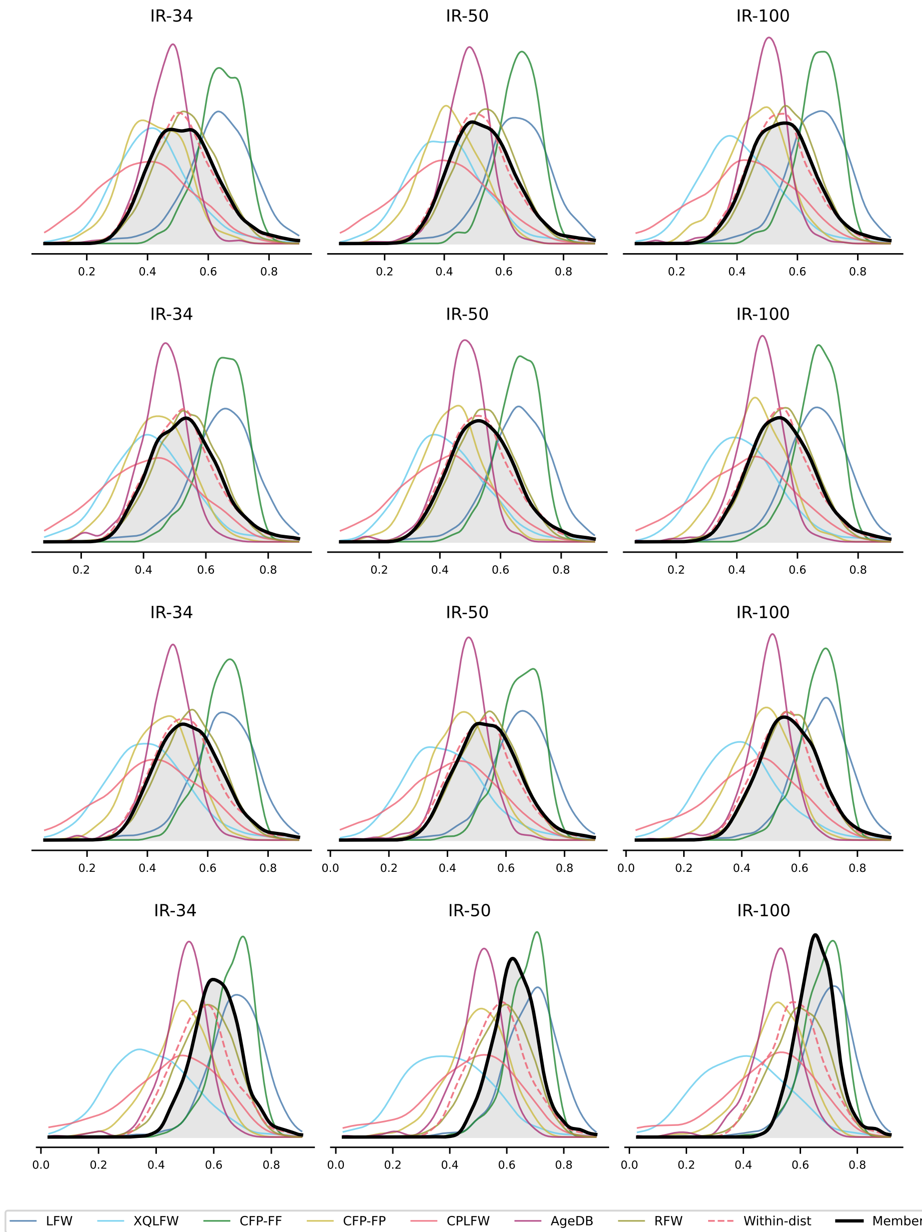


Figure 157. **Vary backbone — ArcFace /** $n_{\text{ids}} = 100K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

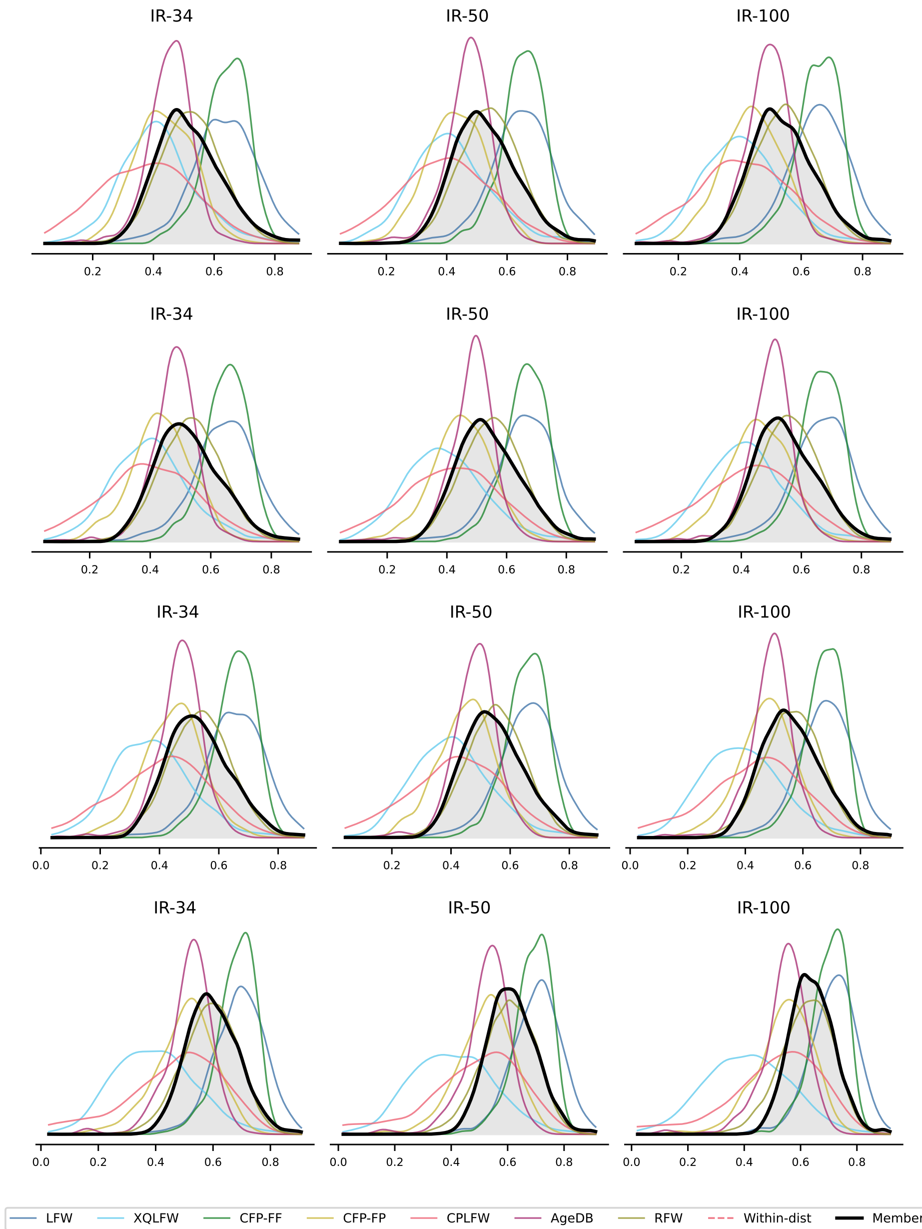


Figure 158. **Vary backbone — ArcFace /** $n_{\text{ids}} = All$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

CosFace

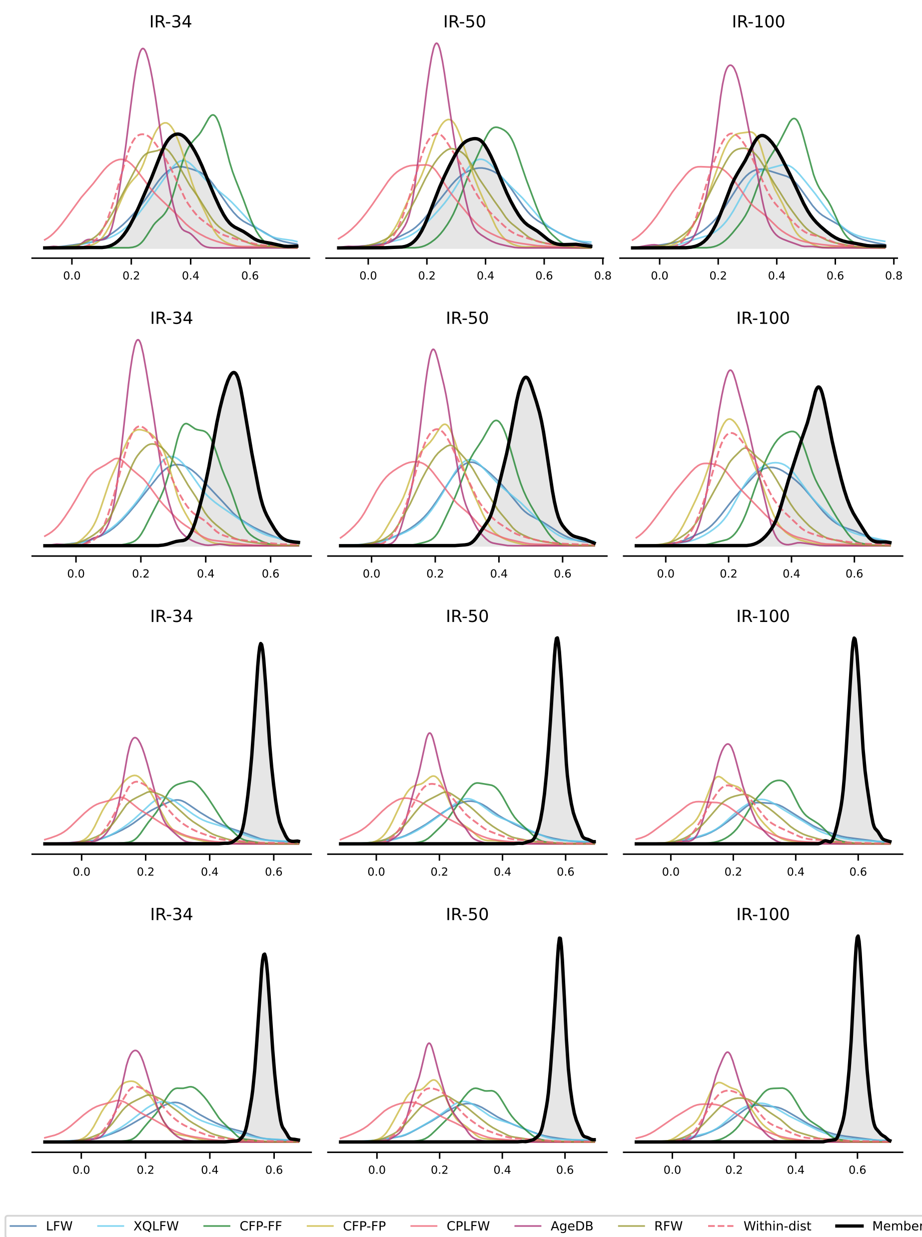


Figure 159. **Vary backbone — CosFace /** $n_{\text{ids}} = 1K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

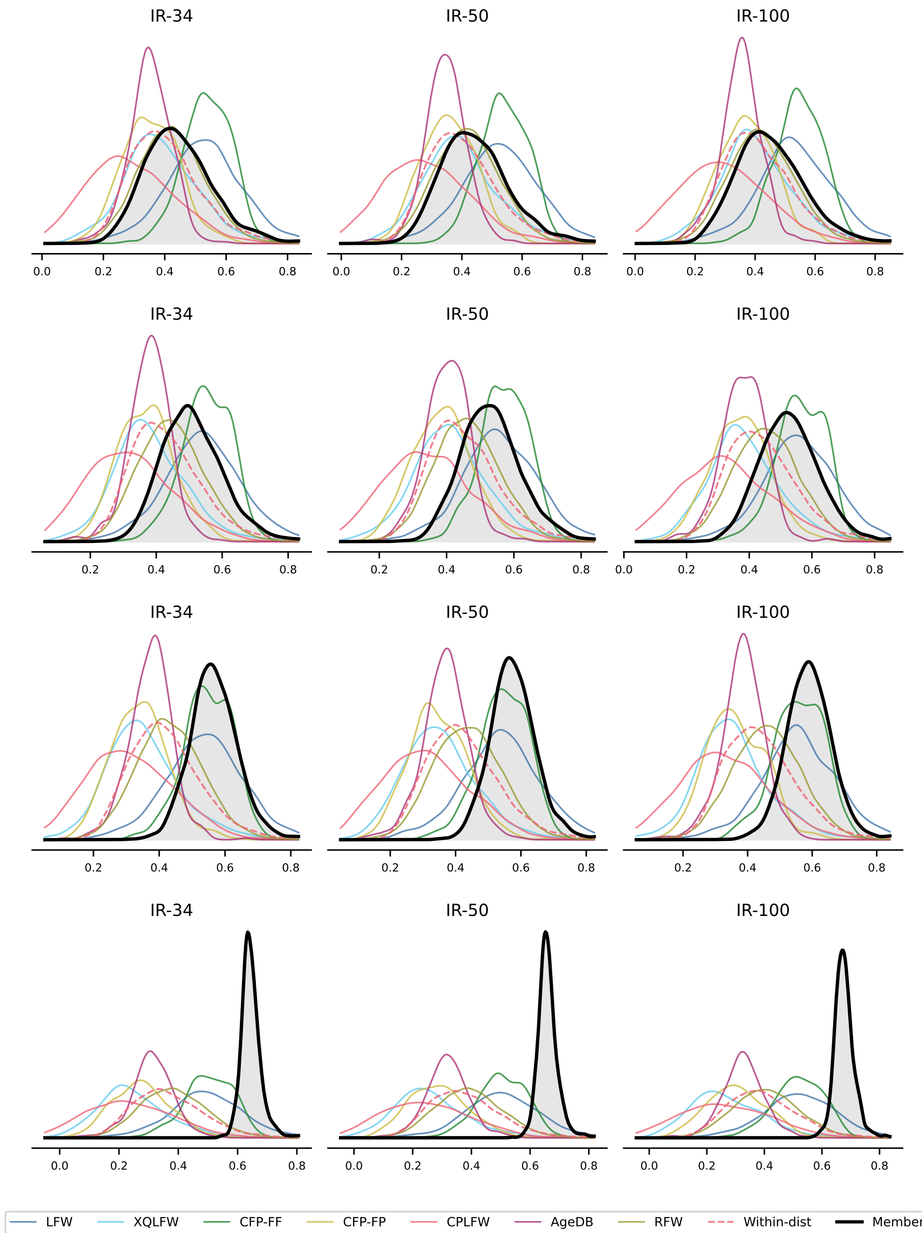


Figure 160. **Vary backbone — CosFace /** $n_{\text{ids}} = 10K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

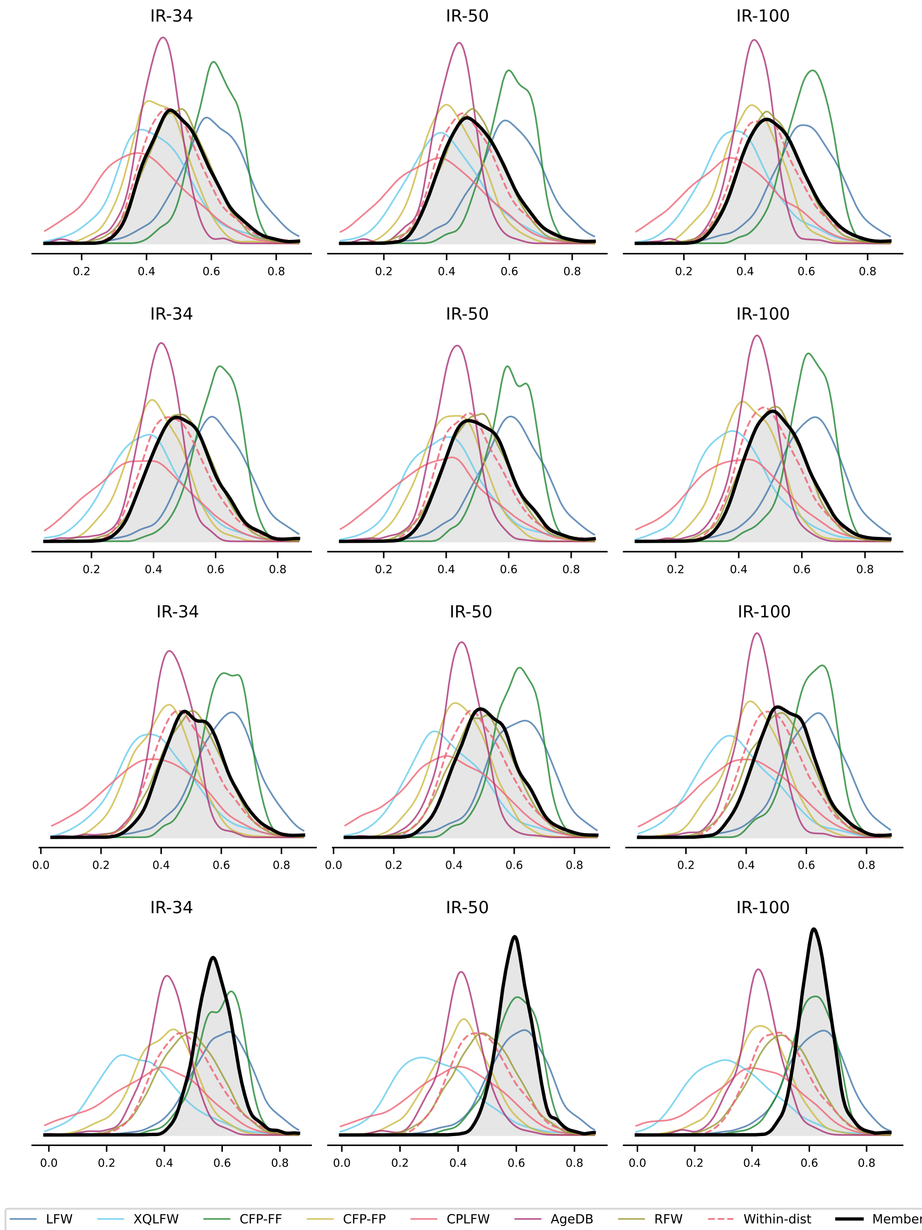


Figure 161. **Vary backbone — CosFace /** $n_{\text{ids}} = 50K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

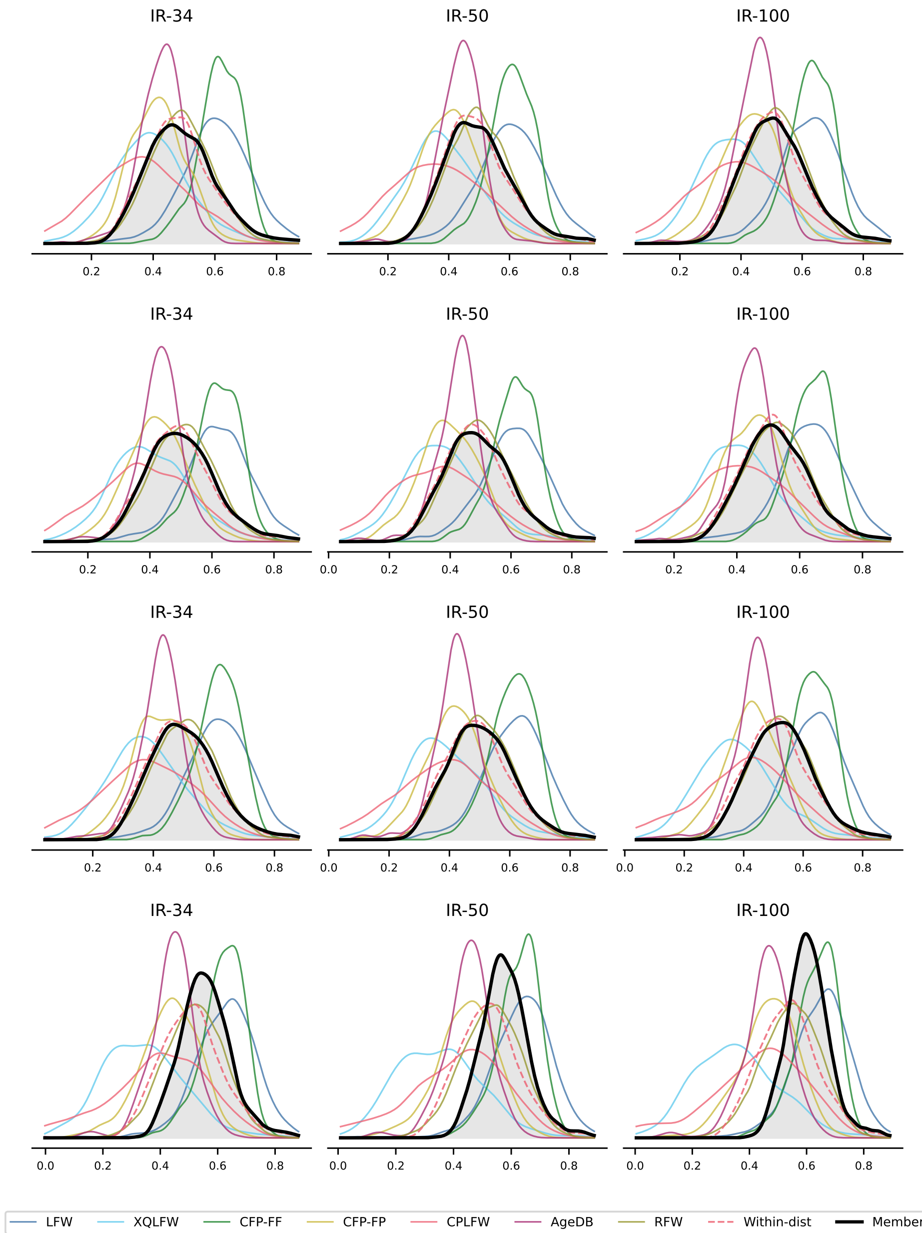


Figure 162. **Vary backbone — CosFace /** $n_{\textbf{ids}} = 100K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

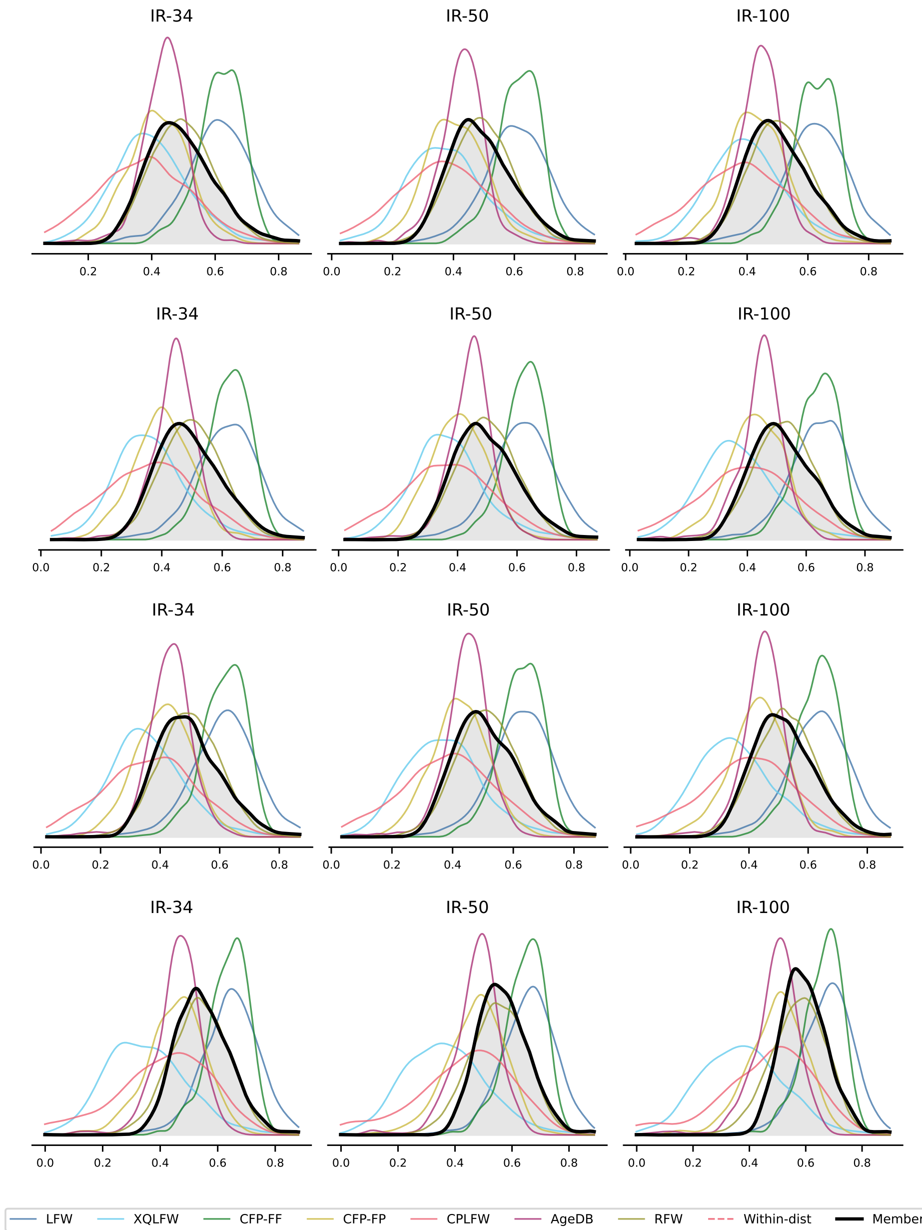


Figure 163. **Vary backbone — CosFace /** $n_{\text{ids}} = All$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

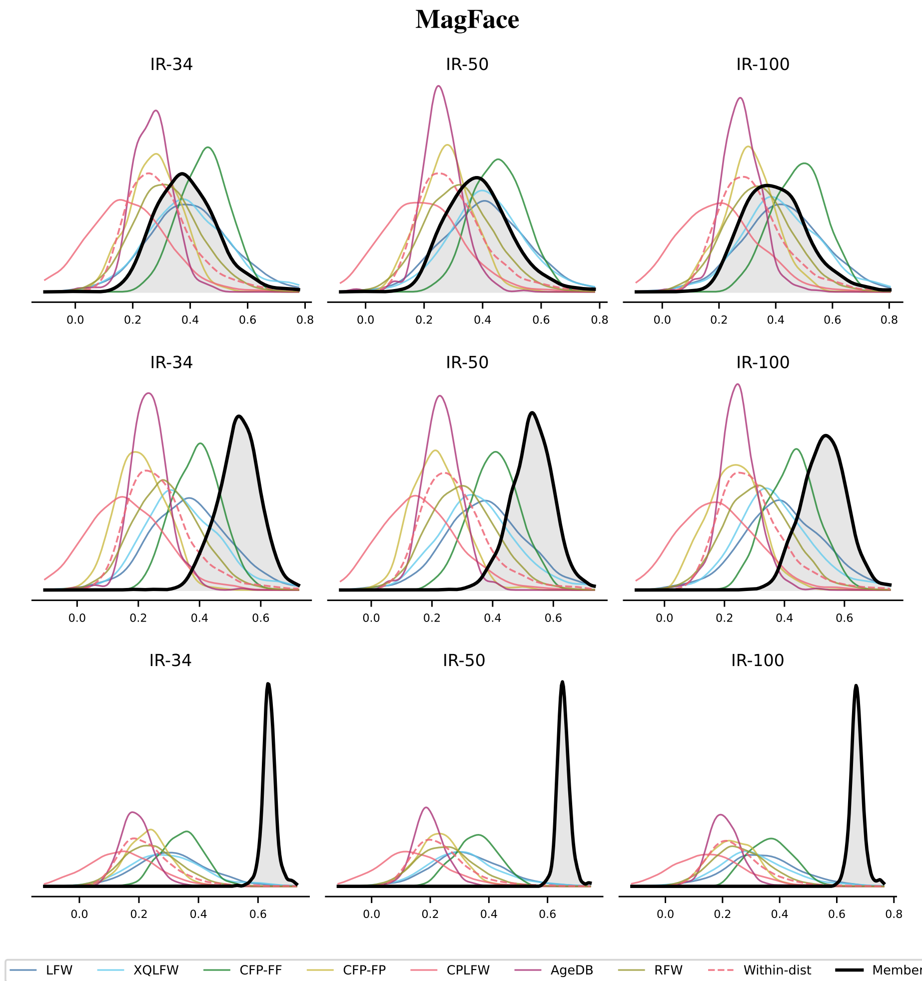


Figure 164. **Vary backbone — MagFace / $n_{\text{ids}} = 1K$.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

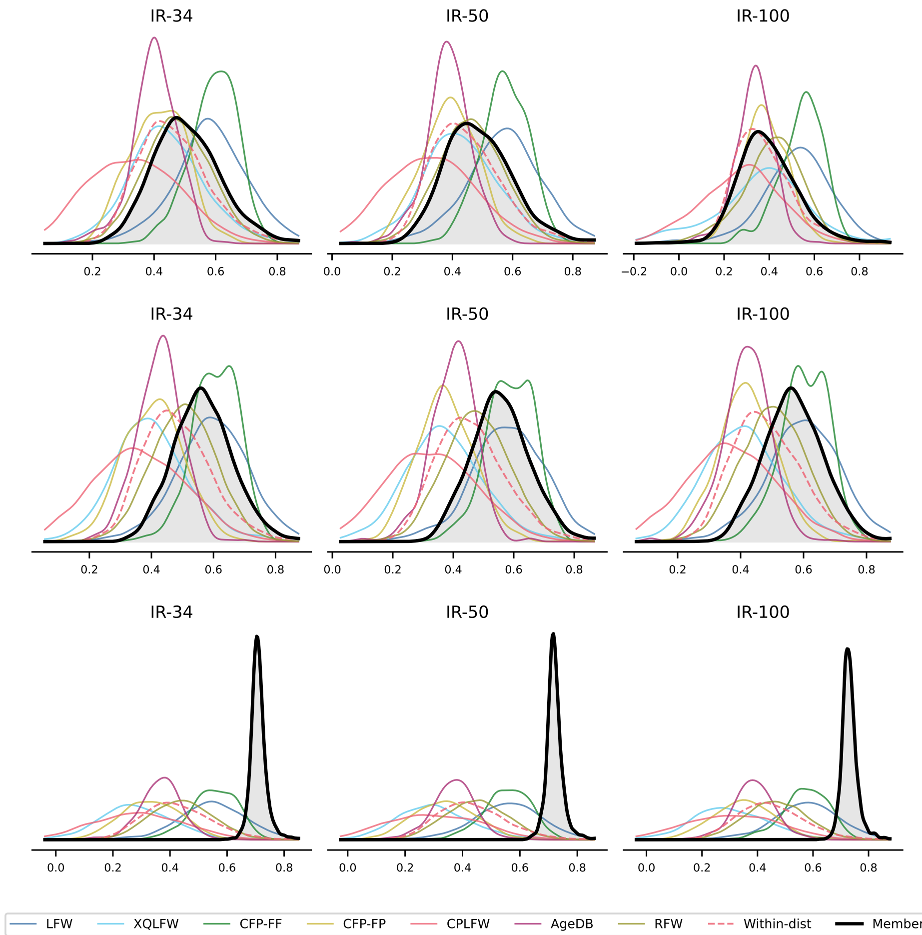


Figure 165. **Vary backbone — MagFace /** $n_{\text{ids}} = 10K$**.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

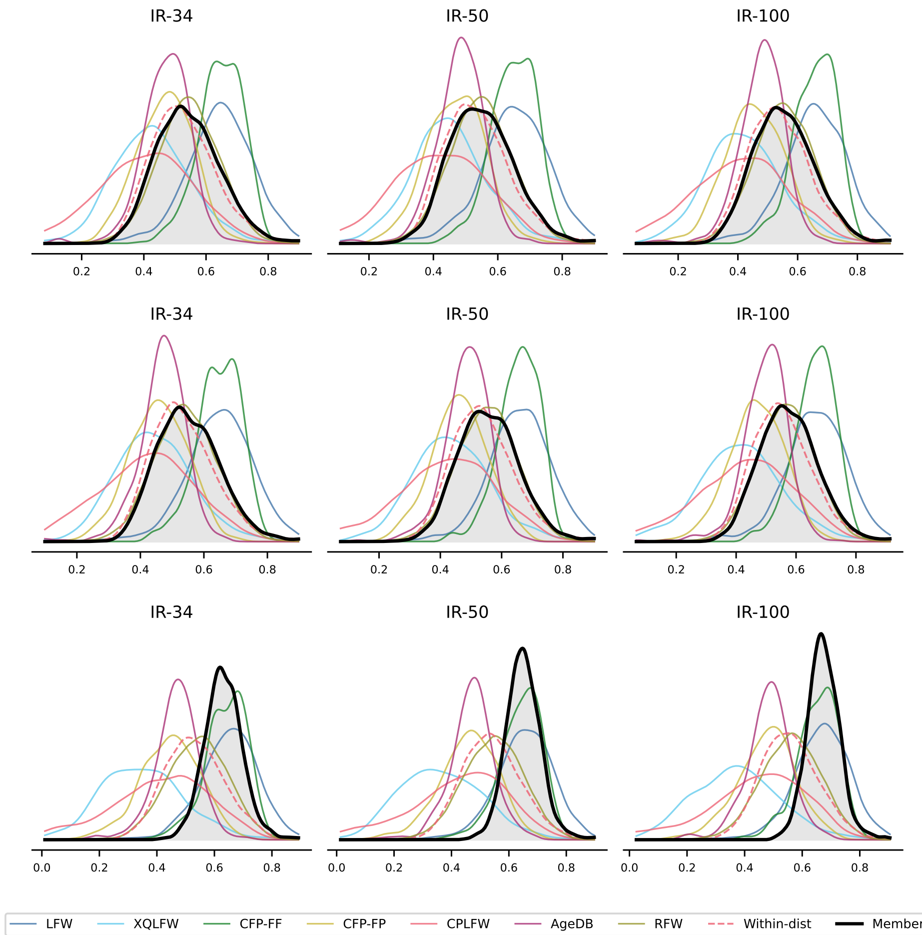


Figure 166. **Vary backbone — MagFace /** $n_{\textbf{ids}} = 50K$**.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

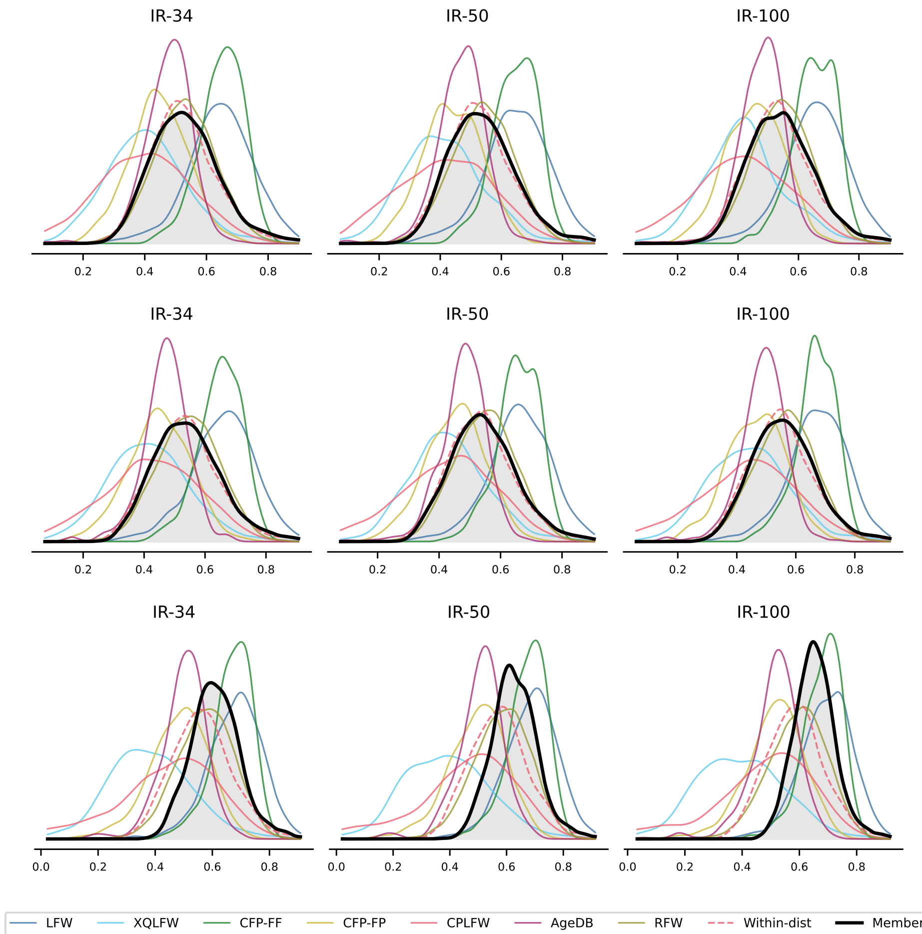


Figure 167. **Vary backbone — MagFace /** $n_{\textbf{ids}} = 100K$**.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

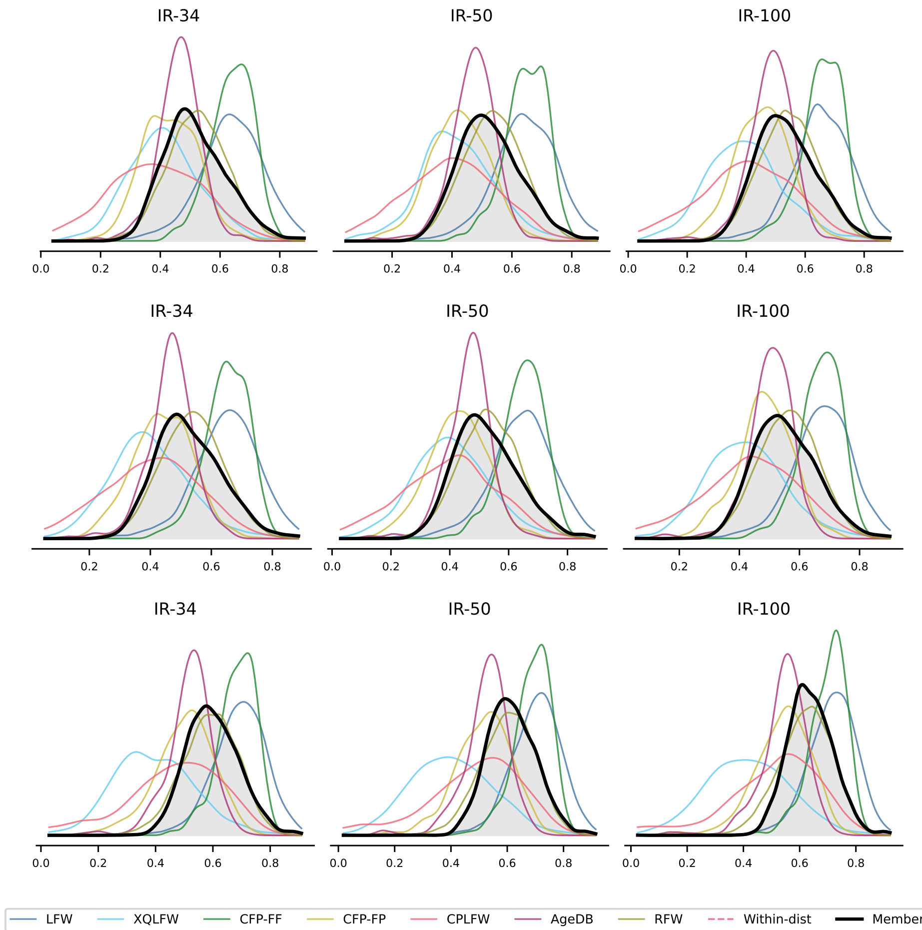


Figure 168. **Vary backbone — MagFace /** $n_{\text{ids}} = All$**.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

5. vMF-κ concentration distributions

5.1. Vary training duration

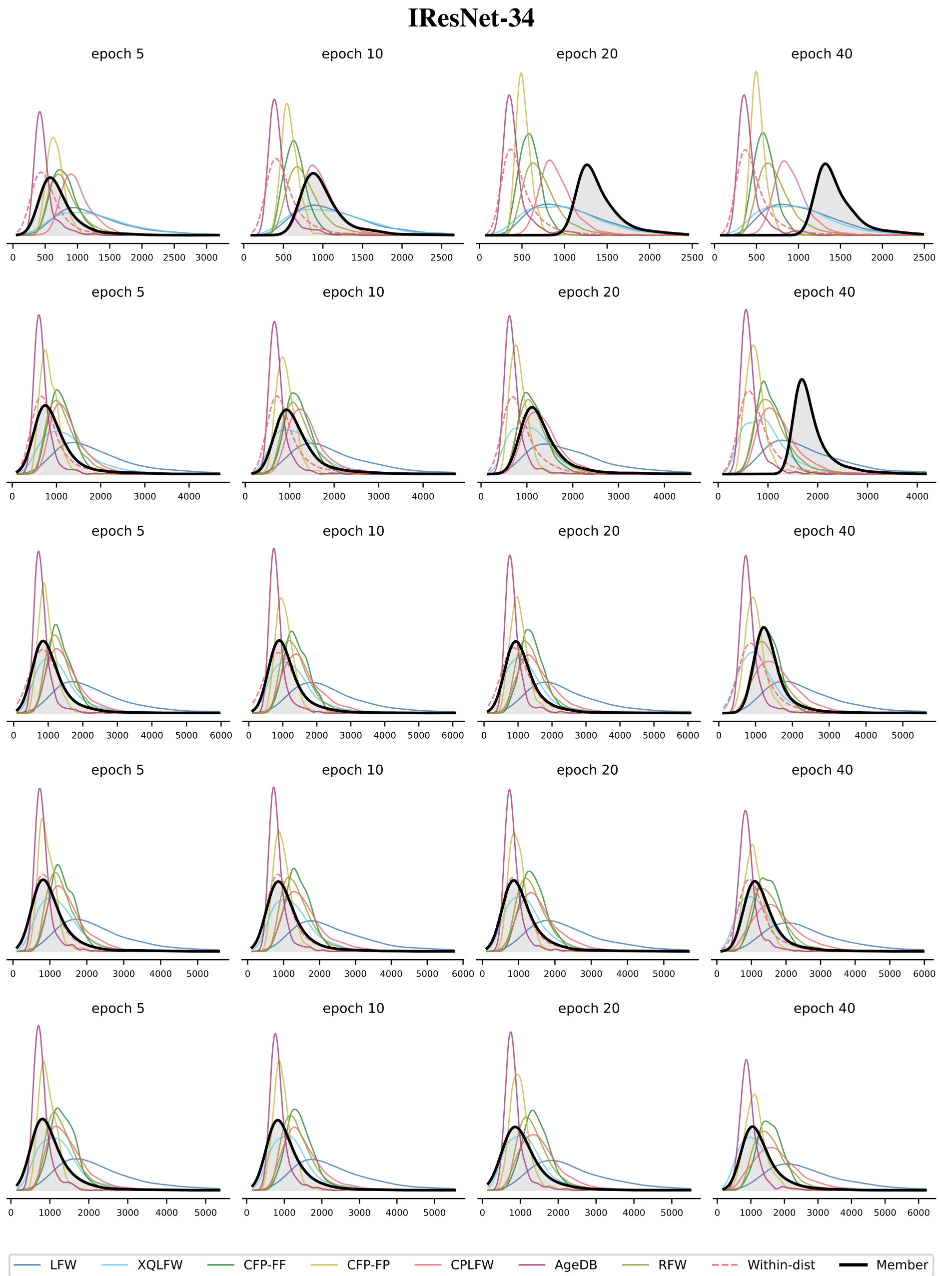


Figure 169. **Vary training duration — IResNet-34 / ArcFace.** Rows: $n_{\text{ids}} \in \{1\text{K},\ 10\text{K},\ 50\text{K},\ 100\text{K},\ \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

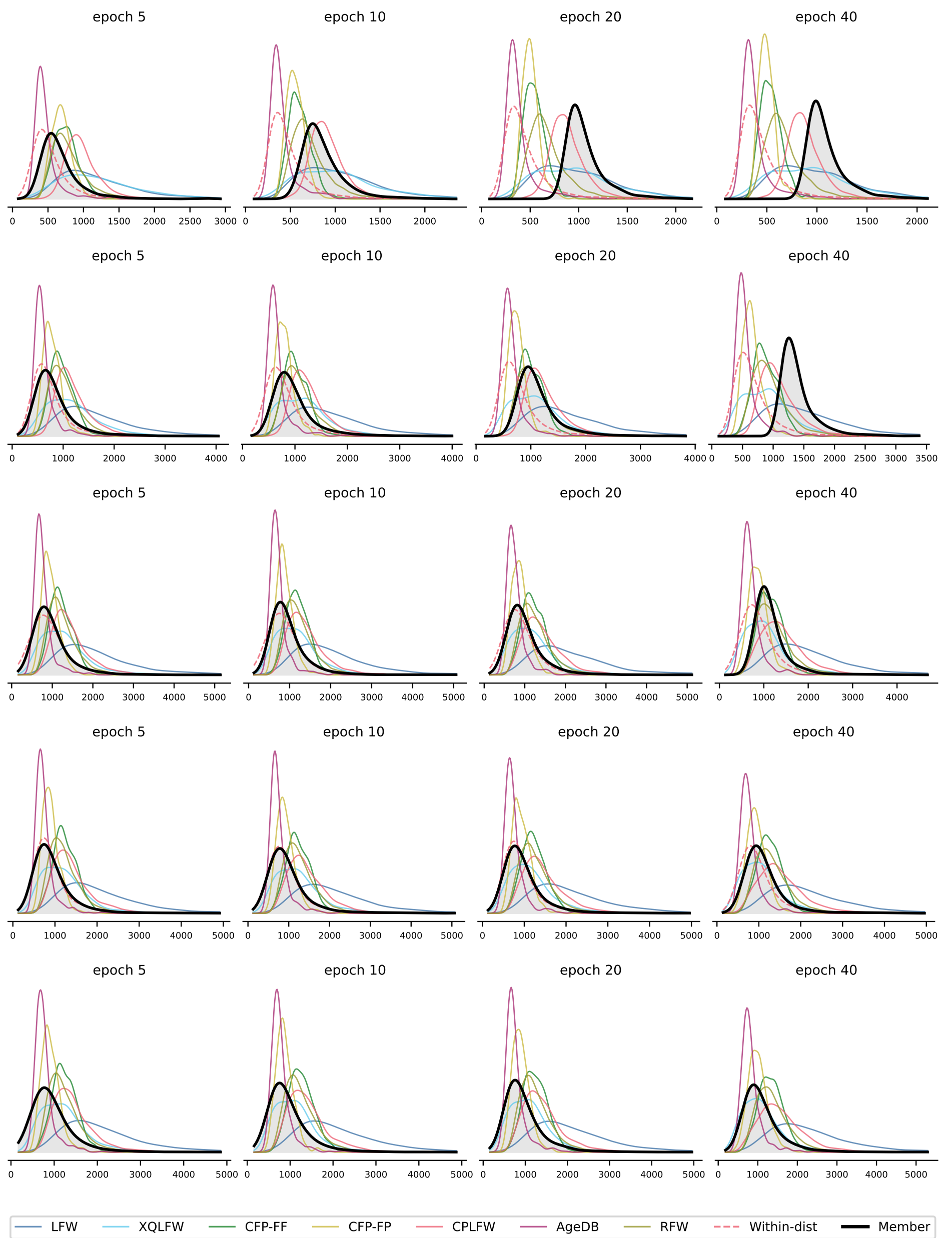


Figure 170. **Vary training duration — IResNet-34 / CosFace.** Rows: $n_{\text{ids}} \in \{1\text{K},\ 10\text{K},\ 50\text{K},\ 100\text{K},\ \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

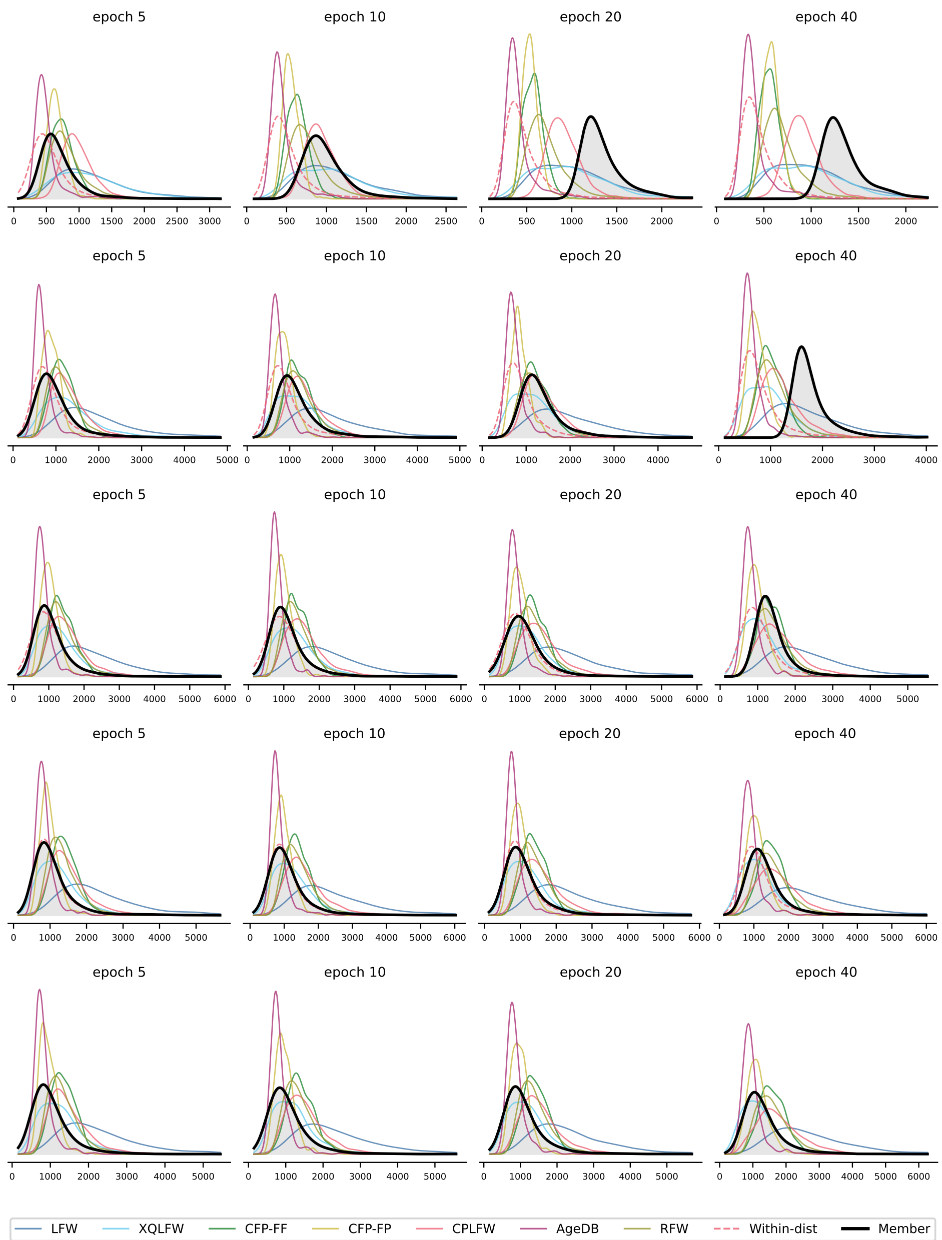


Figure 171. **Vary training duration — IResNet-34 / MagFace.** Rows: $n_{\text{ids}} \in \{1\text{K},\ 10\text{K},\ 50\text{K},\ 100\text{K},\ \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

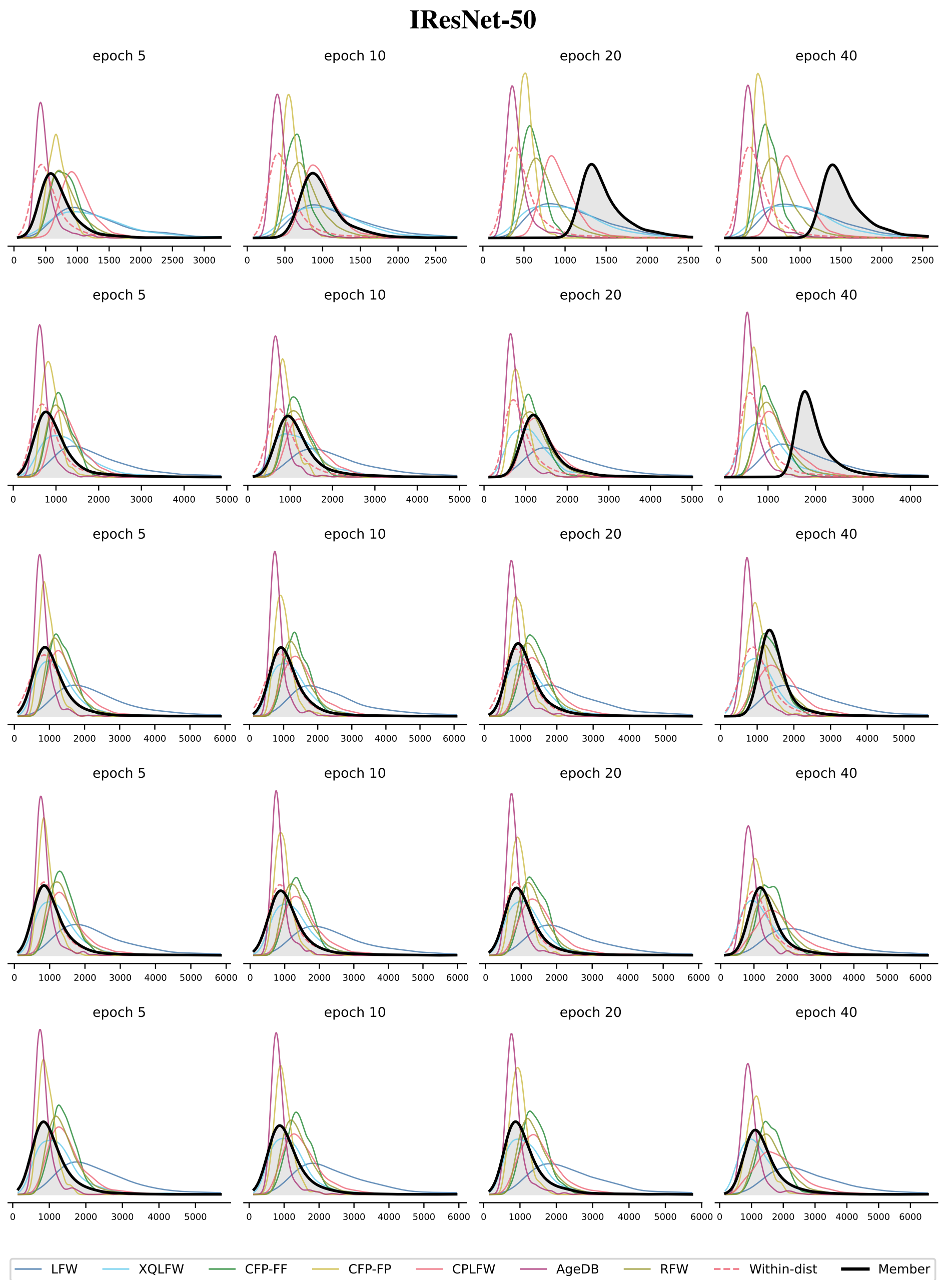


Figure 172. **Vary training duration — IResNet-50 / ArcFace.** Rows: $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

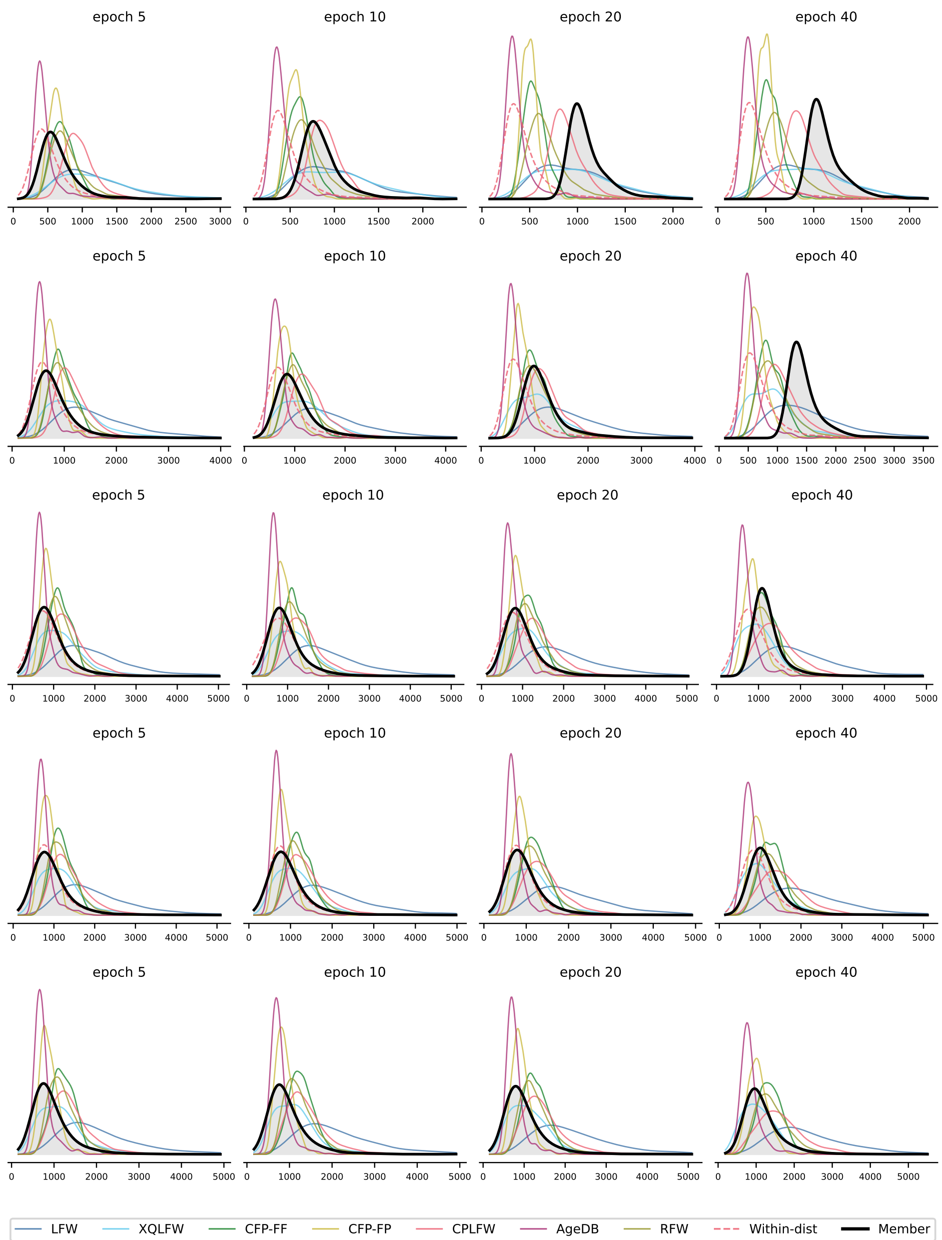


Figure 173. **Vary training duration — IResNet-50 / CosFace.** Rows: $n_{\text{ids}} \in \{1\text{K},\ 10\text{K},\ 50\text{K},\ 100\text{K},\ \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

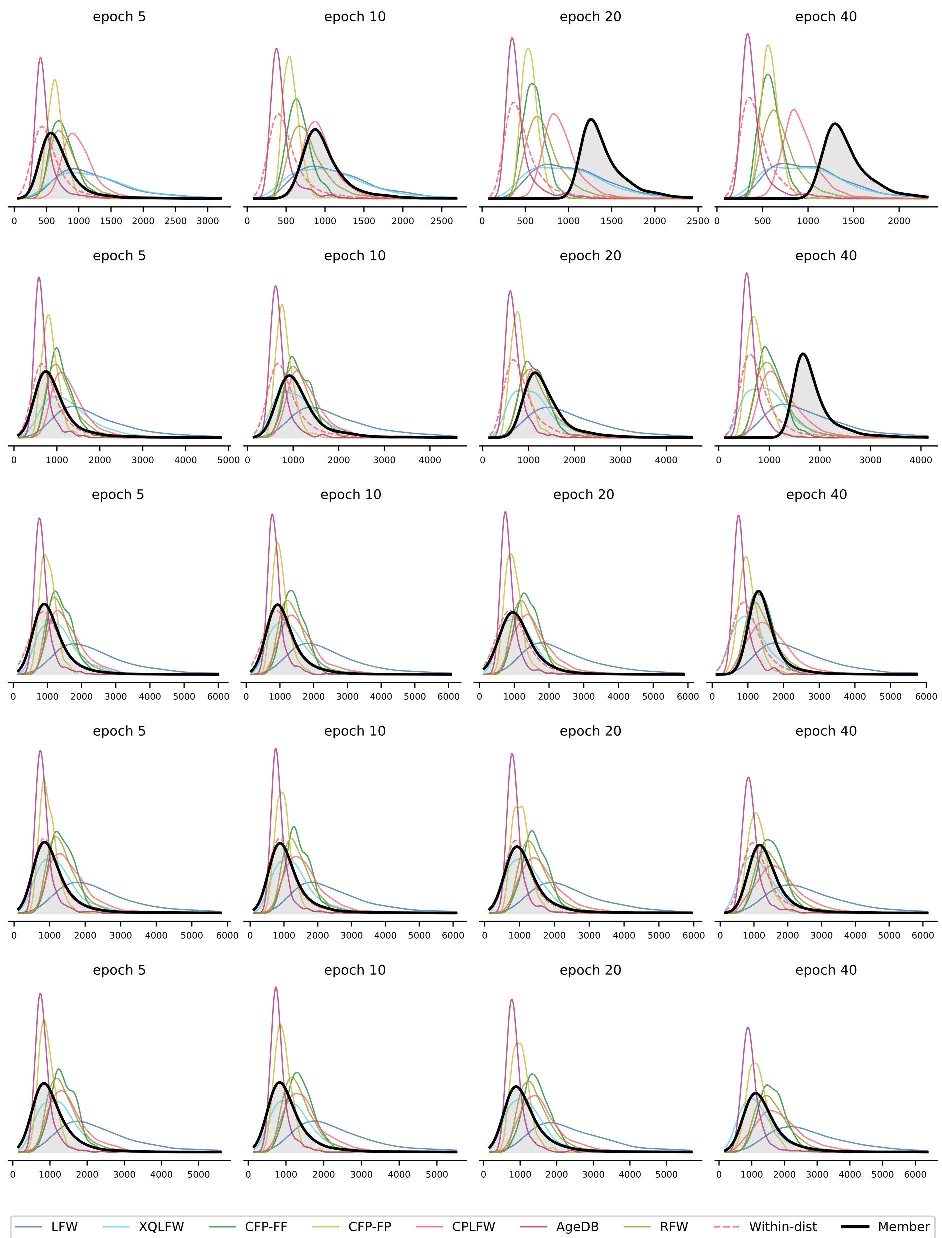


Figure 174. **Vary training duration — IResNet-50 / MagFace.** Rows: $n_{\text{ids}} \in \{1\text{K},\ 10\text{K},\ 50\text{K},\ 100\text{K},\ \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

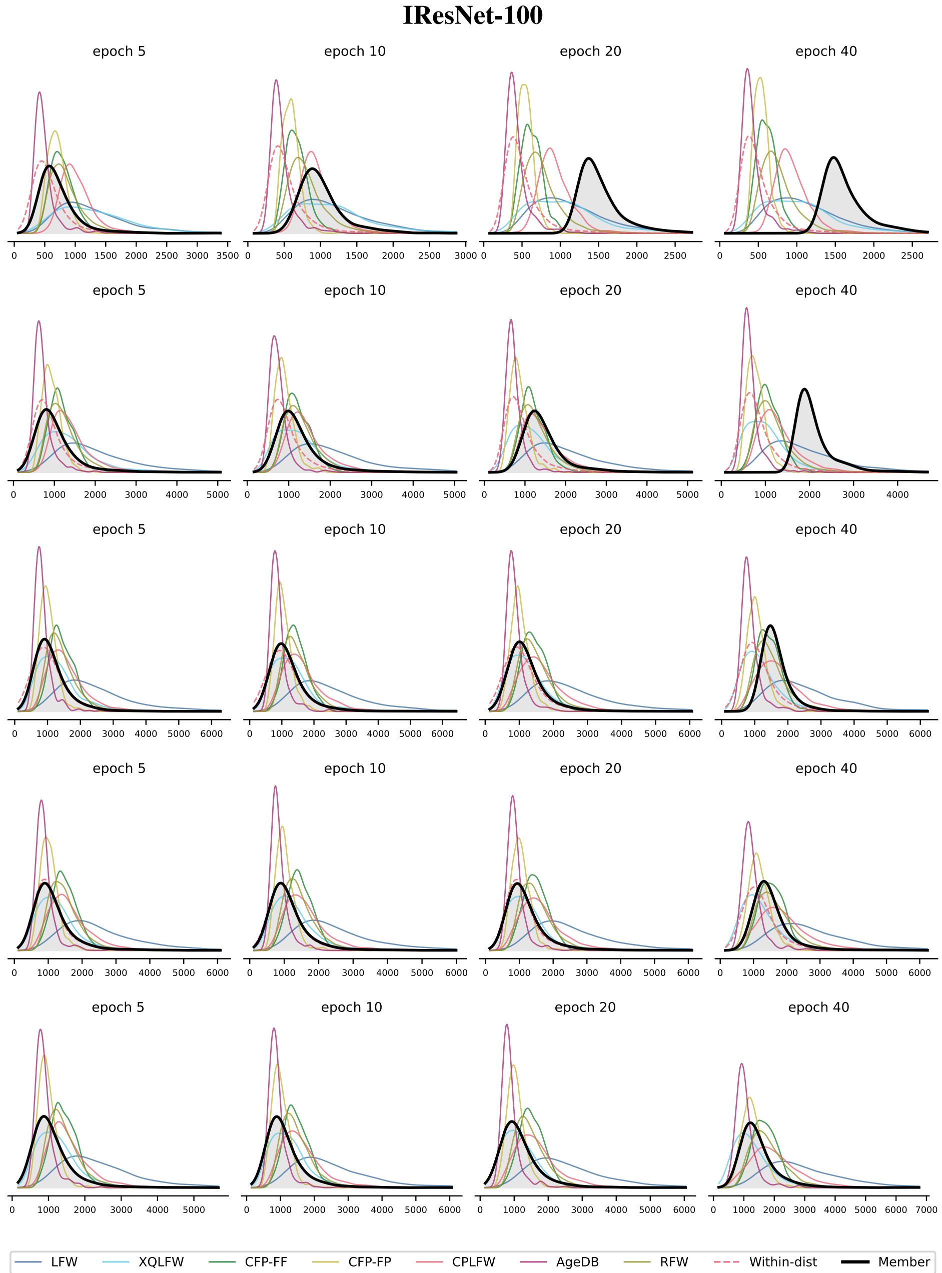


Figure 175. **Vary training duration — IResNet-100 / ArcFace.** Rows: $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

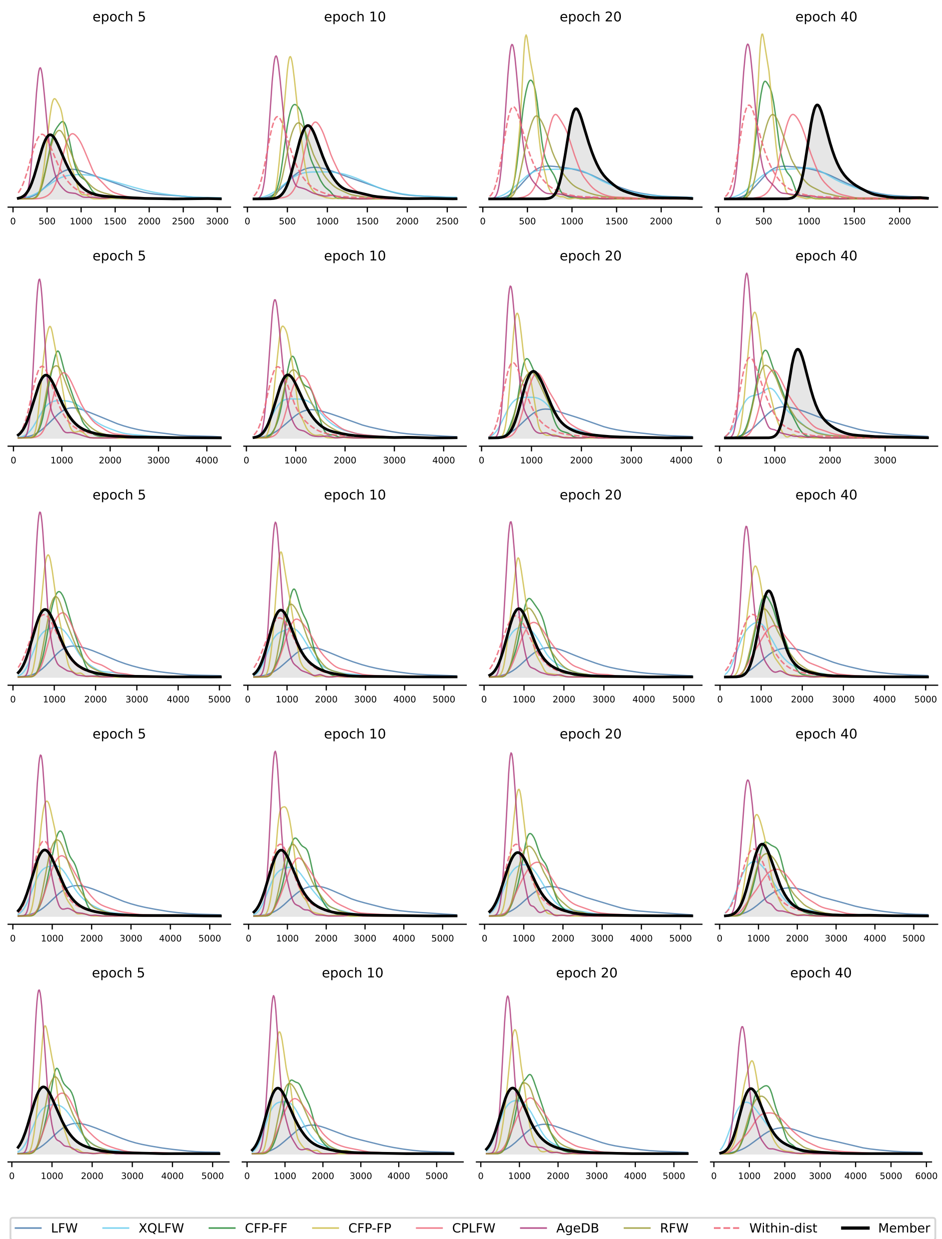


Figure 176. **Vary training duration — IResNet-100 / CosFace.** Rows: $n_{\text{ids}} \in \{1\text{K},\ 10\text{K},\ 50\text{K},\ 100\text{K},\ \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

5.2. Vary training-set size

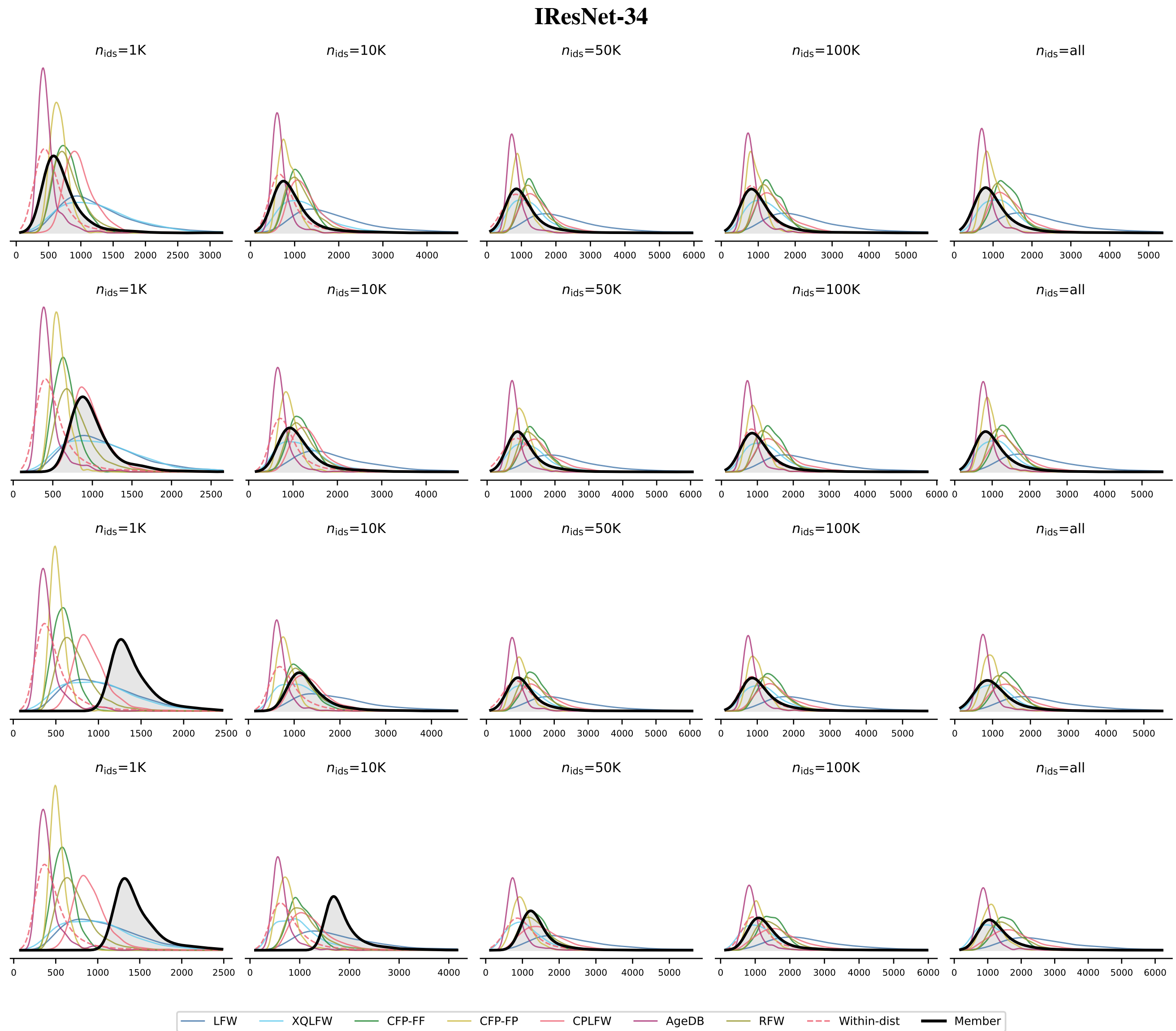


Figure 177. **Vary training-set size — IResNet-34 / ArcFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

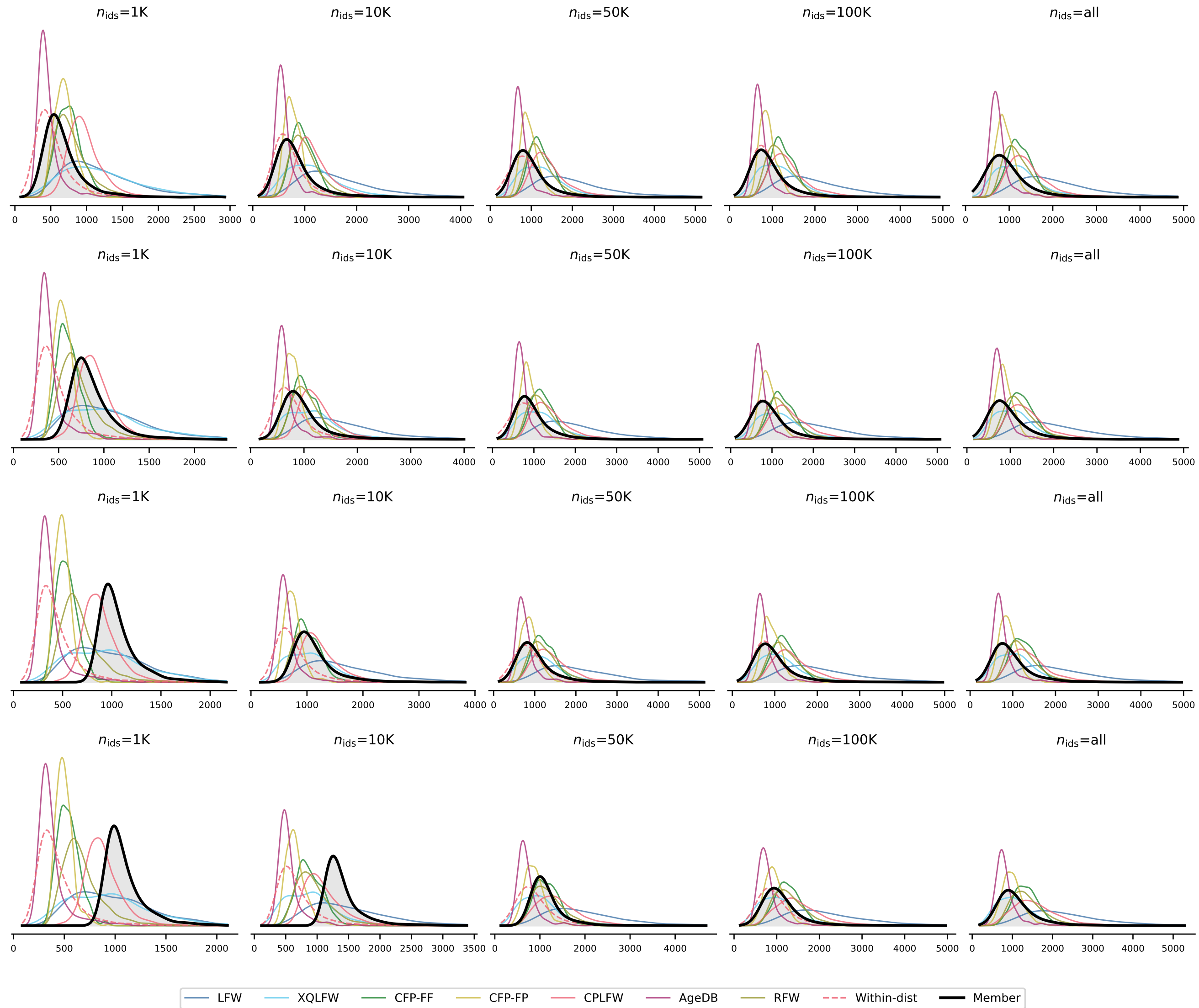


Figure 178. **Vary training-set size — IResNet-34 / CosFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

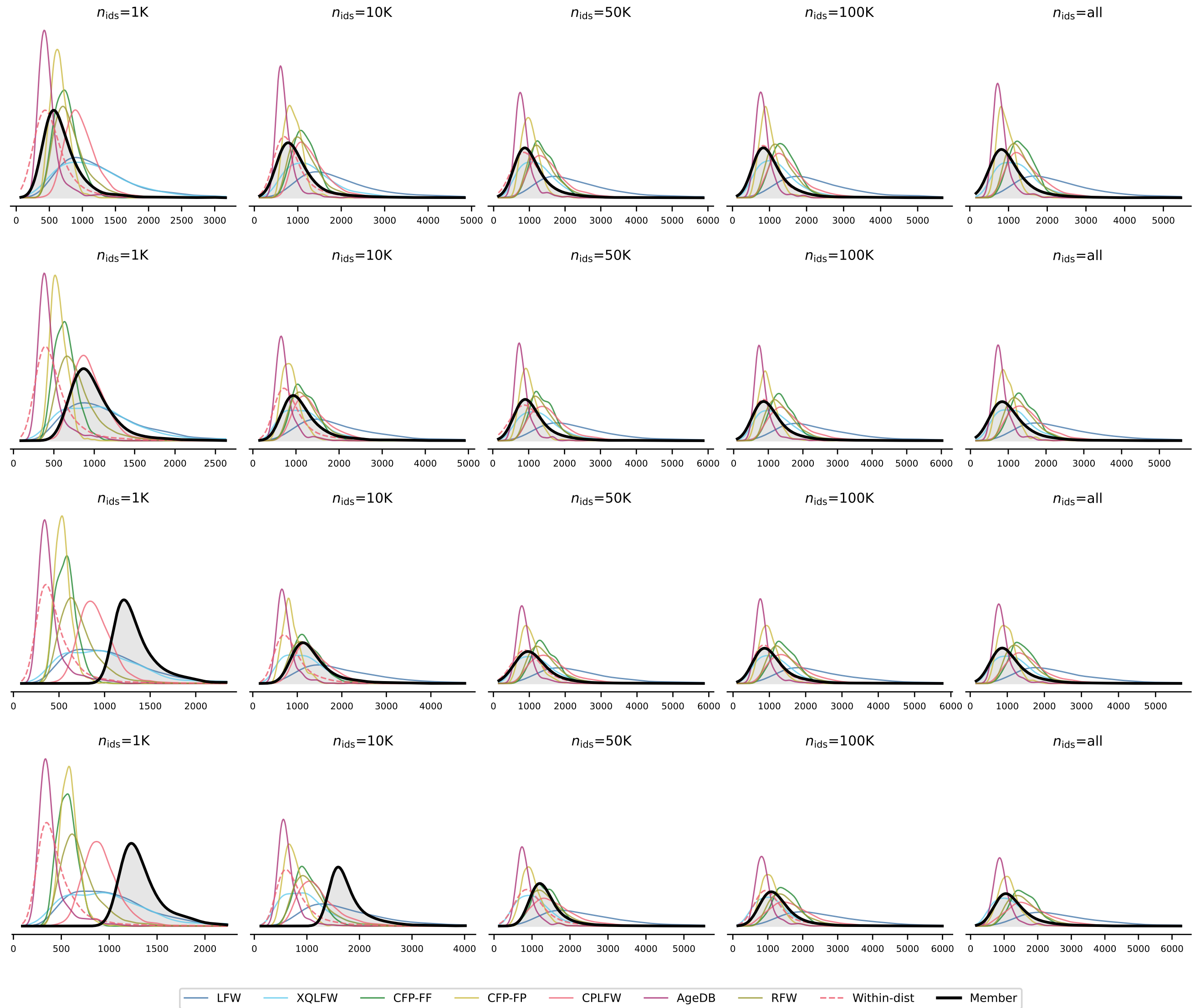


Figure 179. **Vary training-set size — IResNet-34 / MagFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

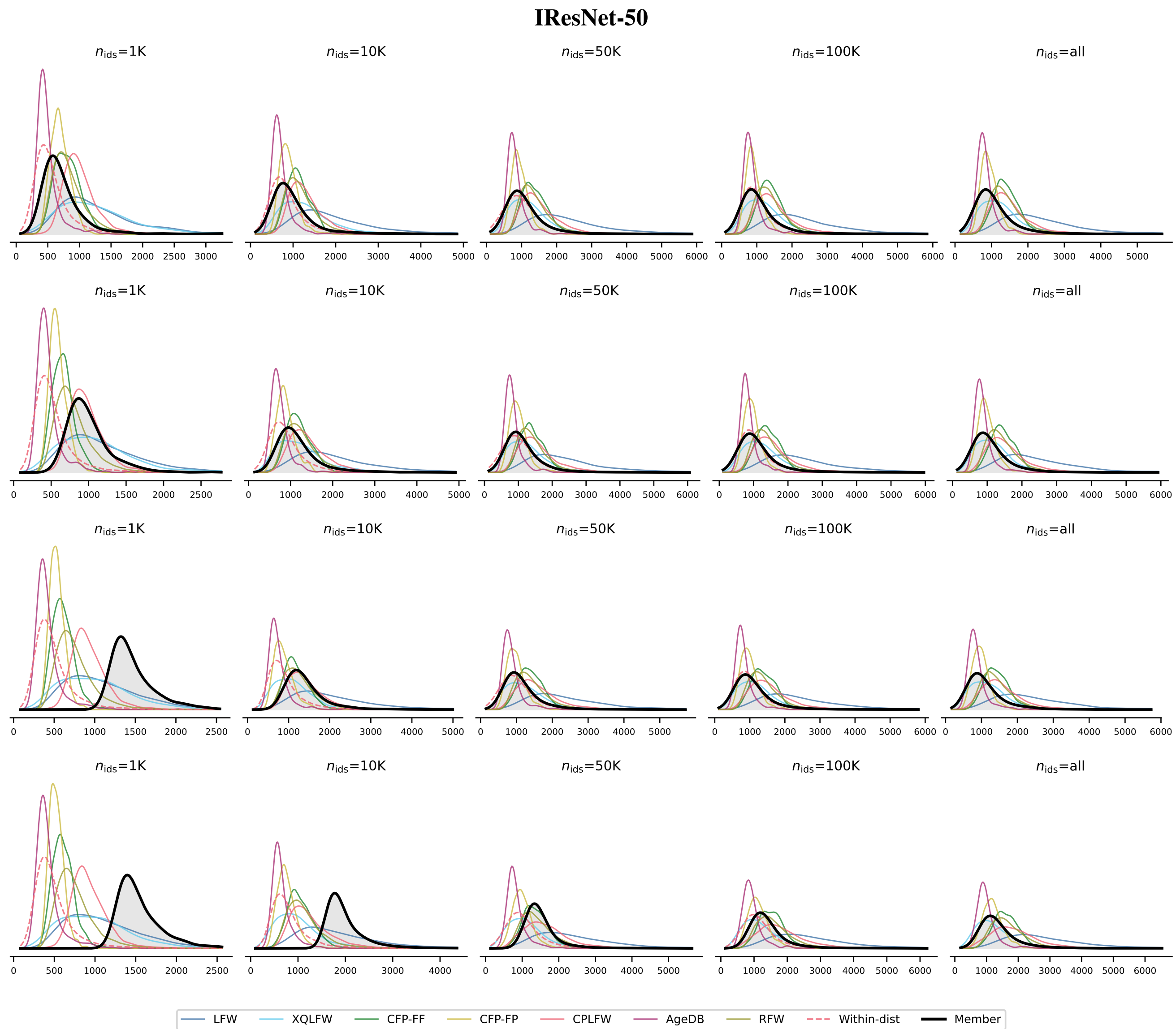


Figure 180. **Vary training-set size — IResNet-50 / ArcFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

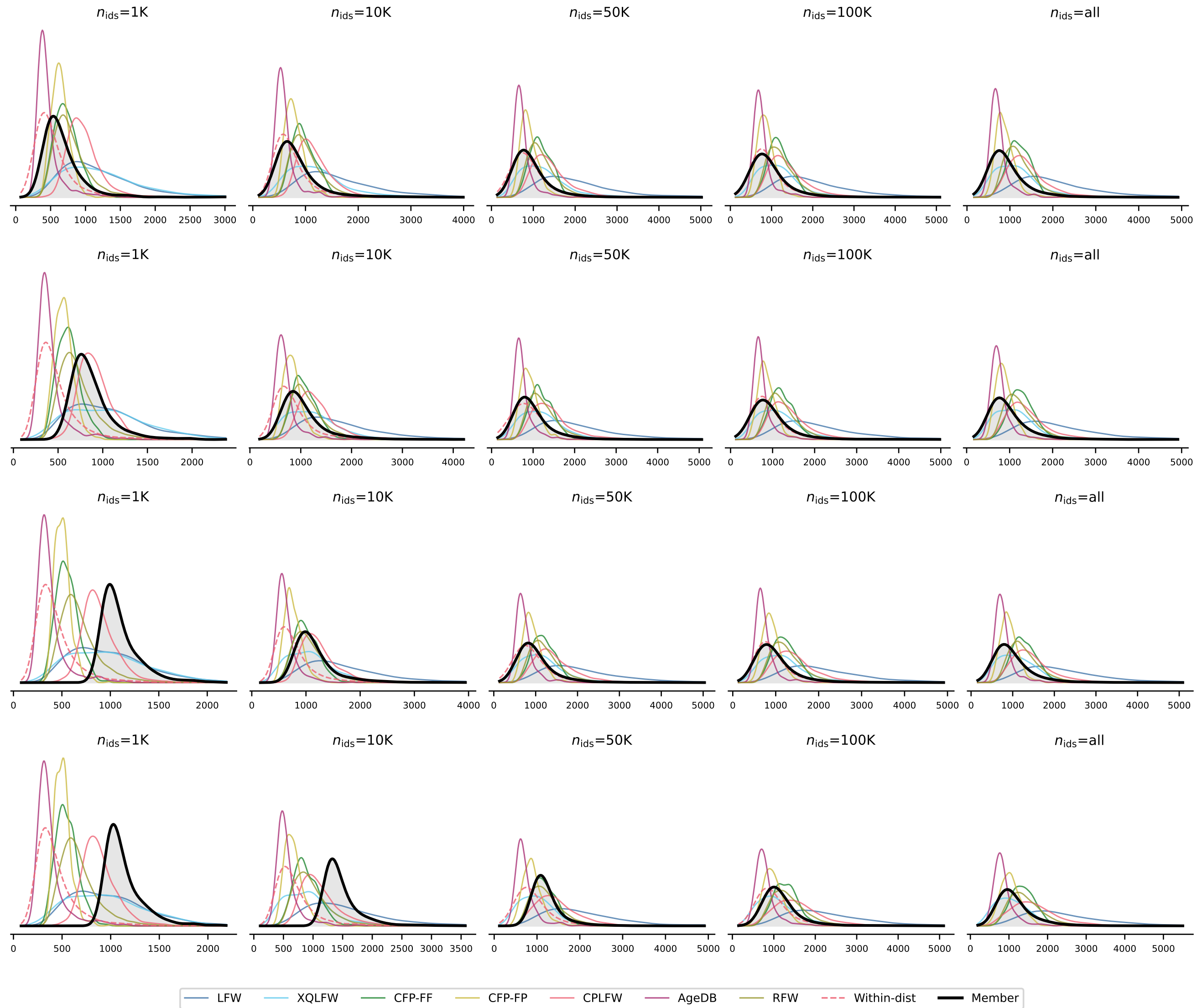


Figure 181. **Vary training-set size — IResNet-50 / CosFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

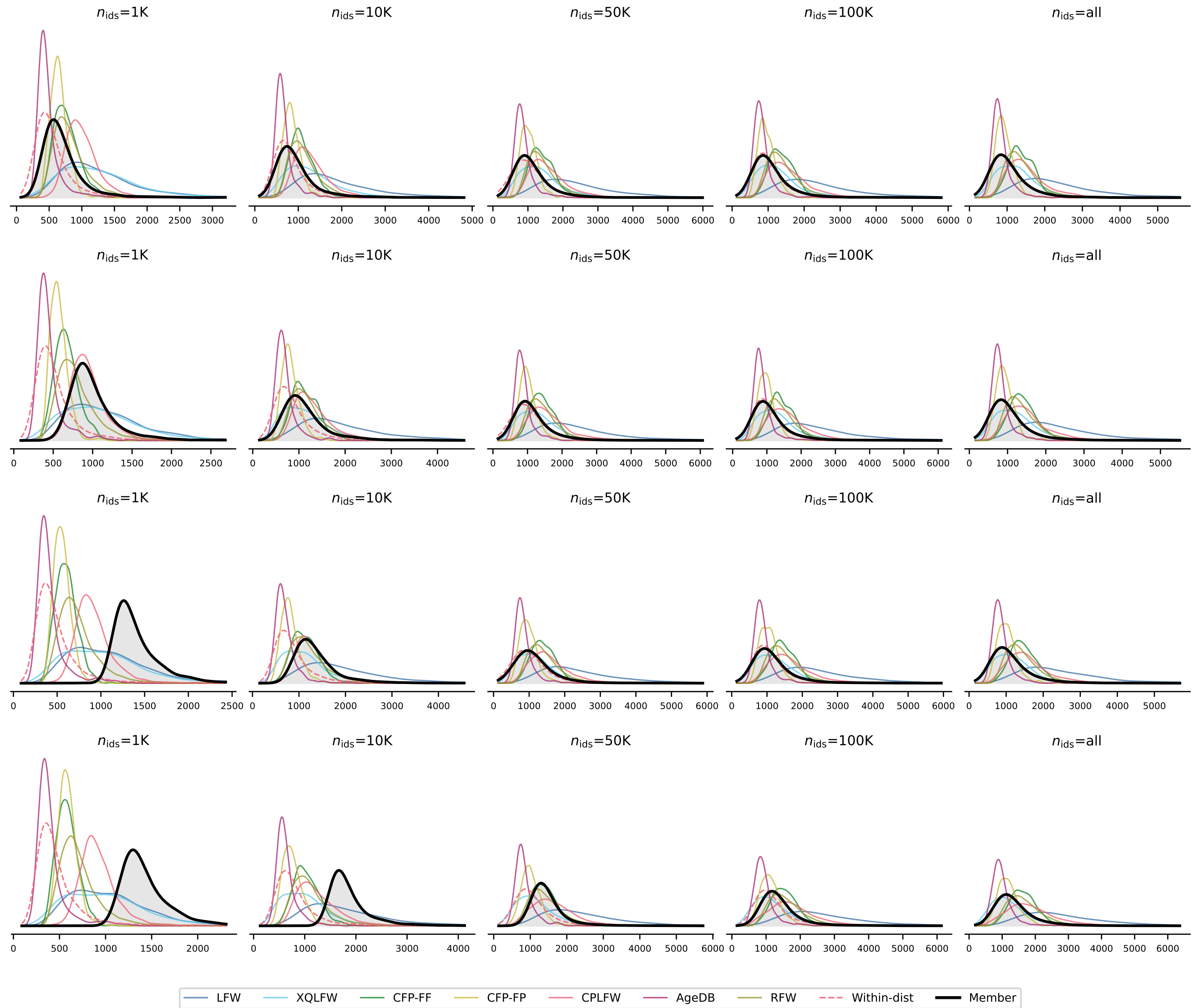


Figure 182. **Vary training-set size — IResNet-50 / MagFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

IResNet-100

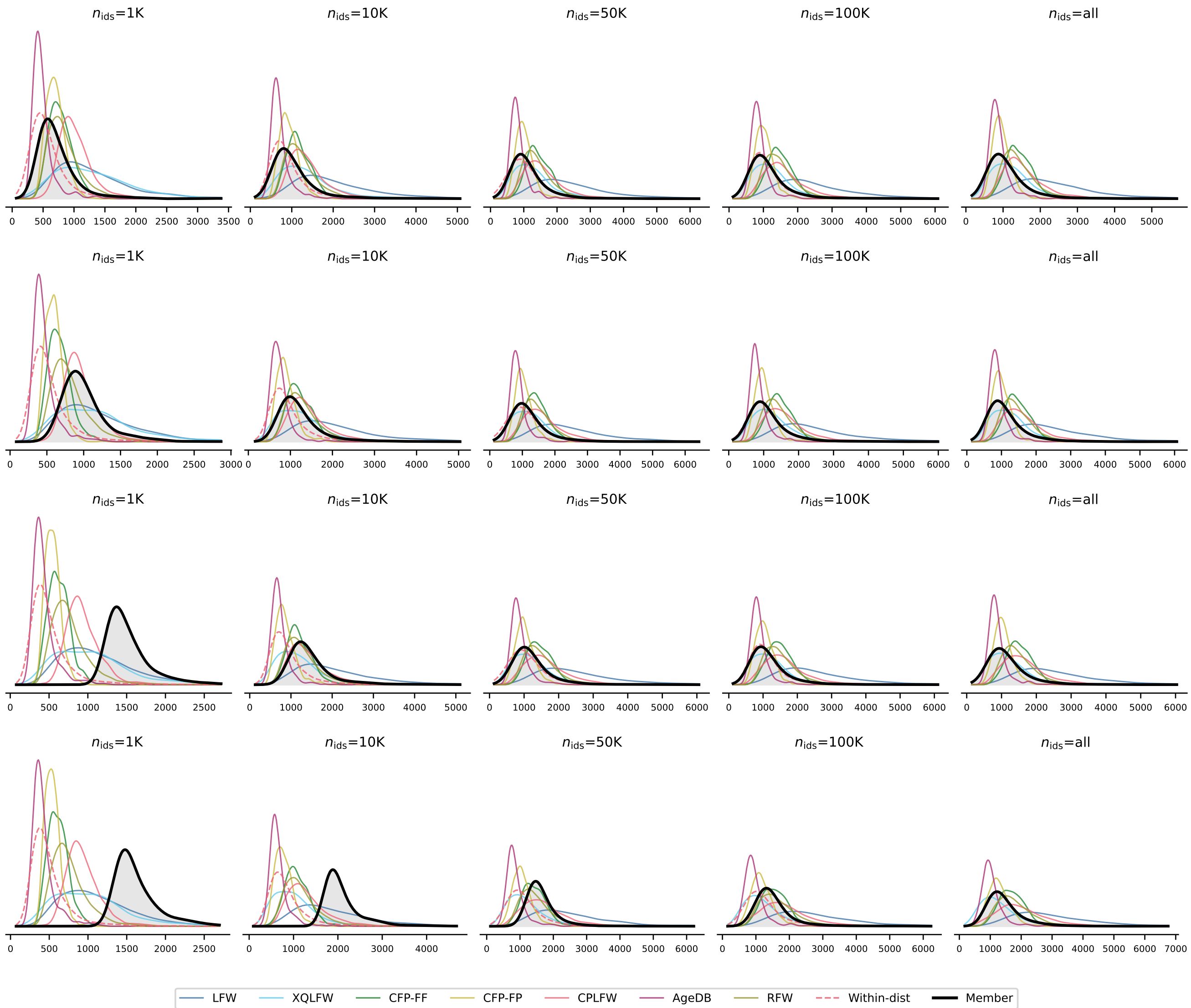


Figure 183. **Vary training-set size — IResNet-100 / ArcFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

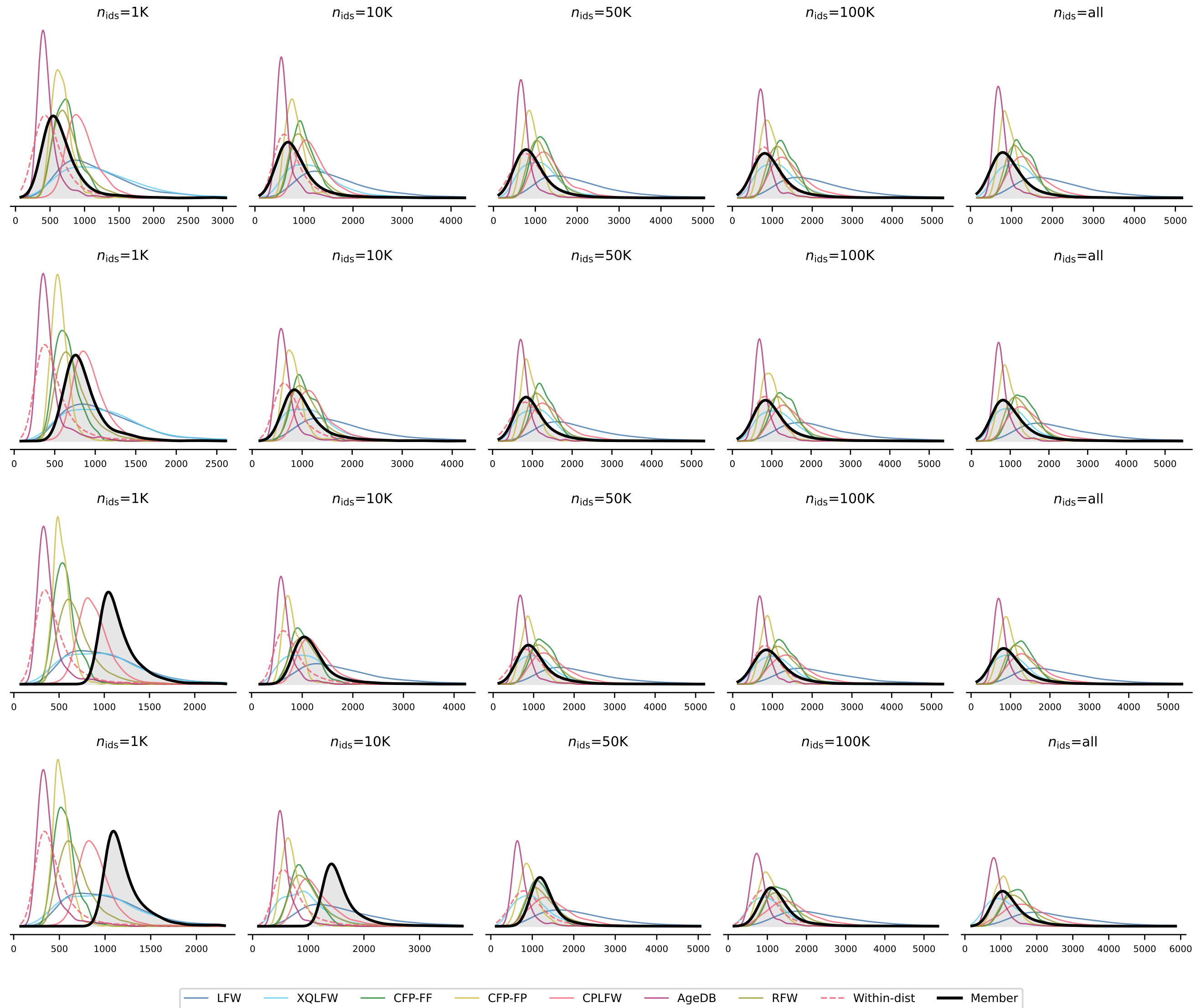


Figure 184. **Vary training-set size — IResNet-100 / CosFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

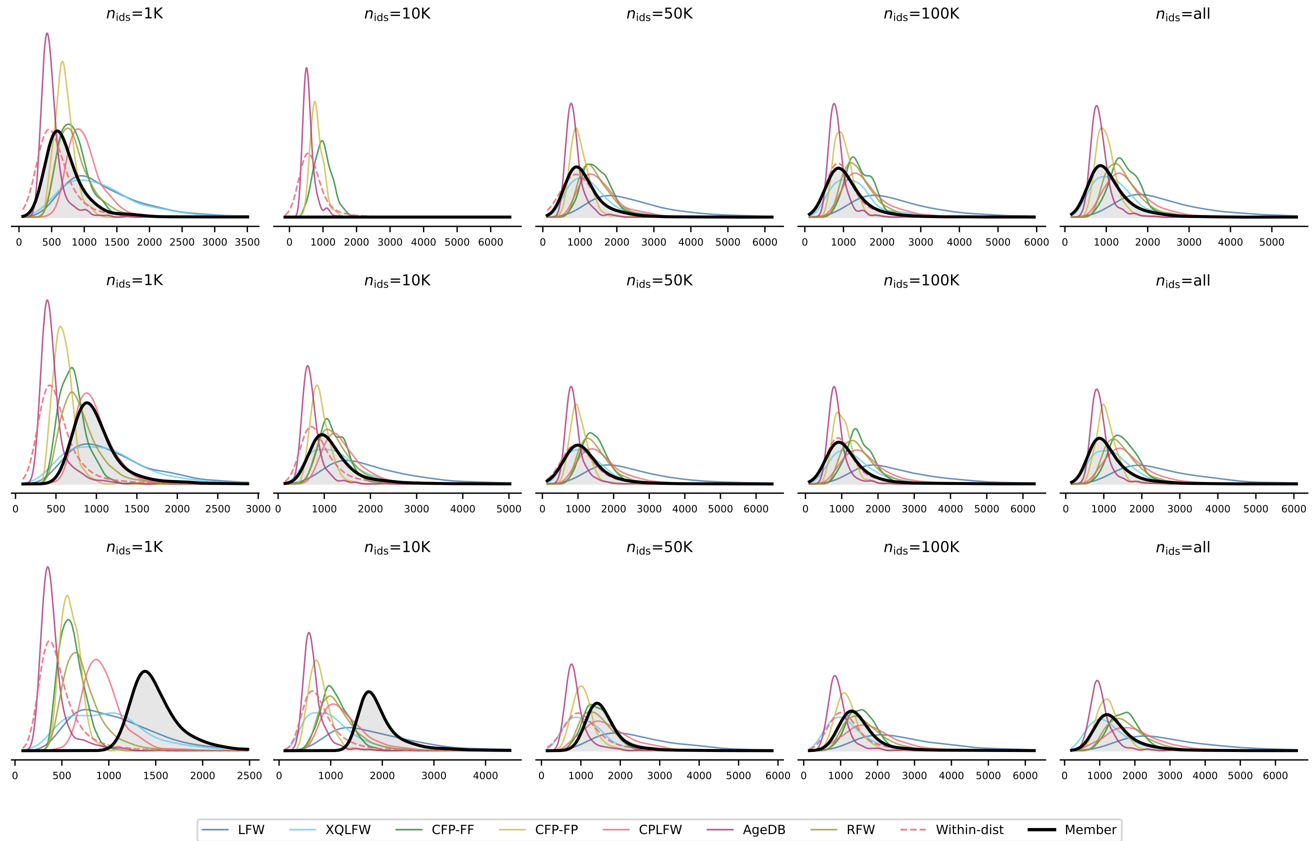


Figure 185. **Vary training-set size — IResNet-100 / MagFace.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

5.3. Vary loss head

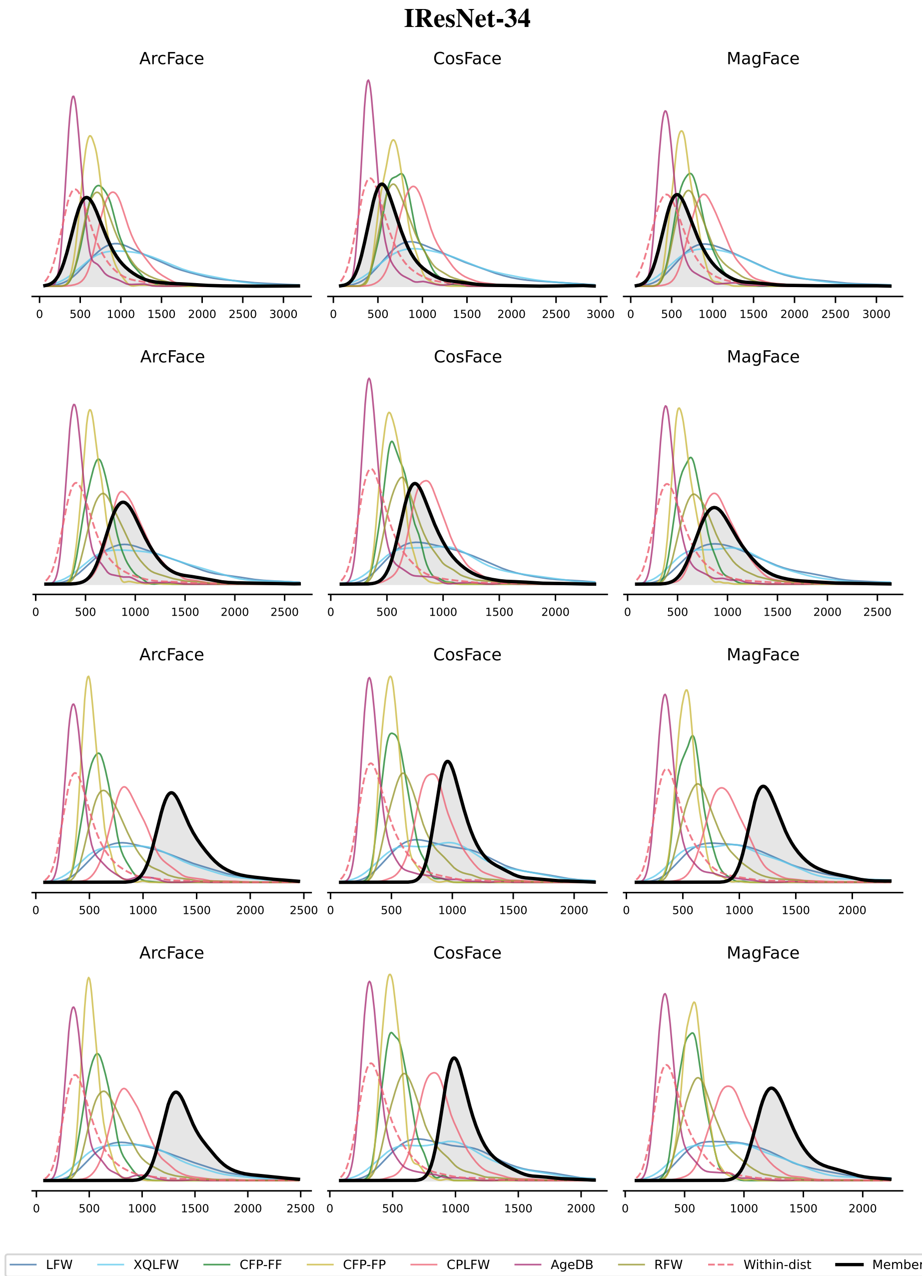


Figure 186. **Vary loss head — IResNet-34 / $n_{\text{ids}} = 1K$.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

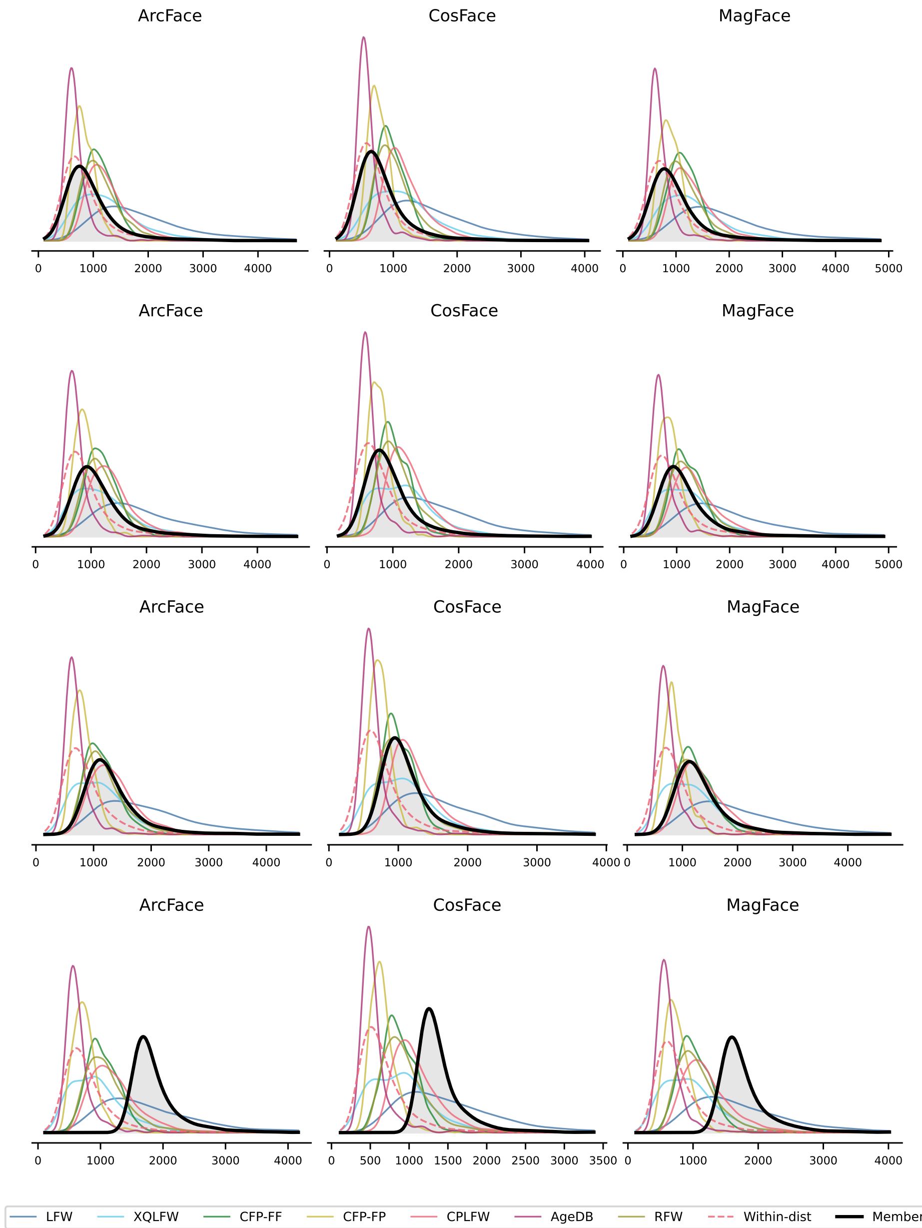


Figure 187. **Vary loss head — IResNet-34 / $n_{\text{ids}} = 10K$.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

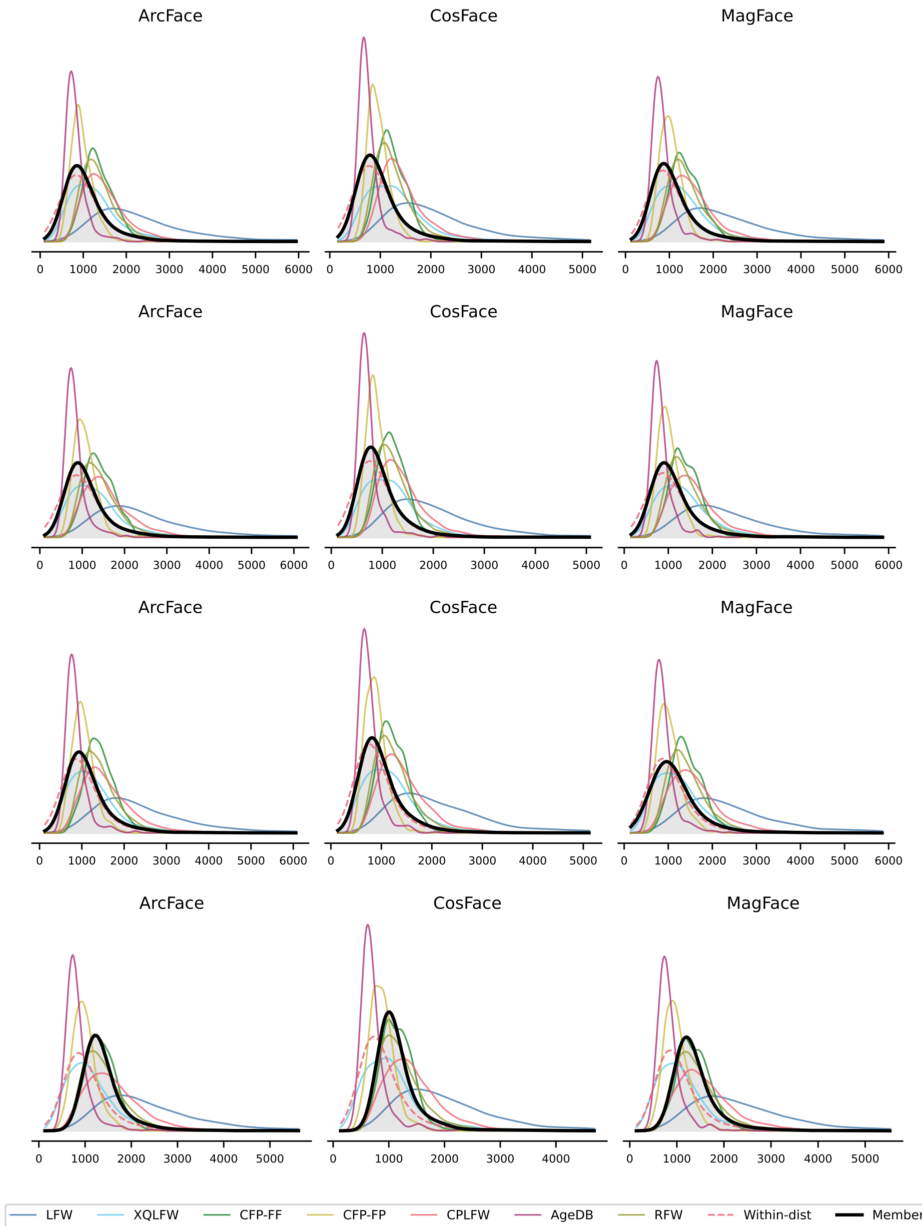


Figure 188. **Vary loss head — IResNet-34 / $n_{\text{ids}} = 50K$.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

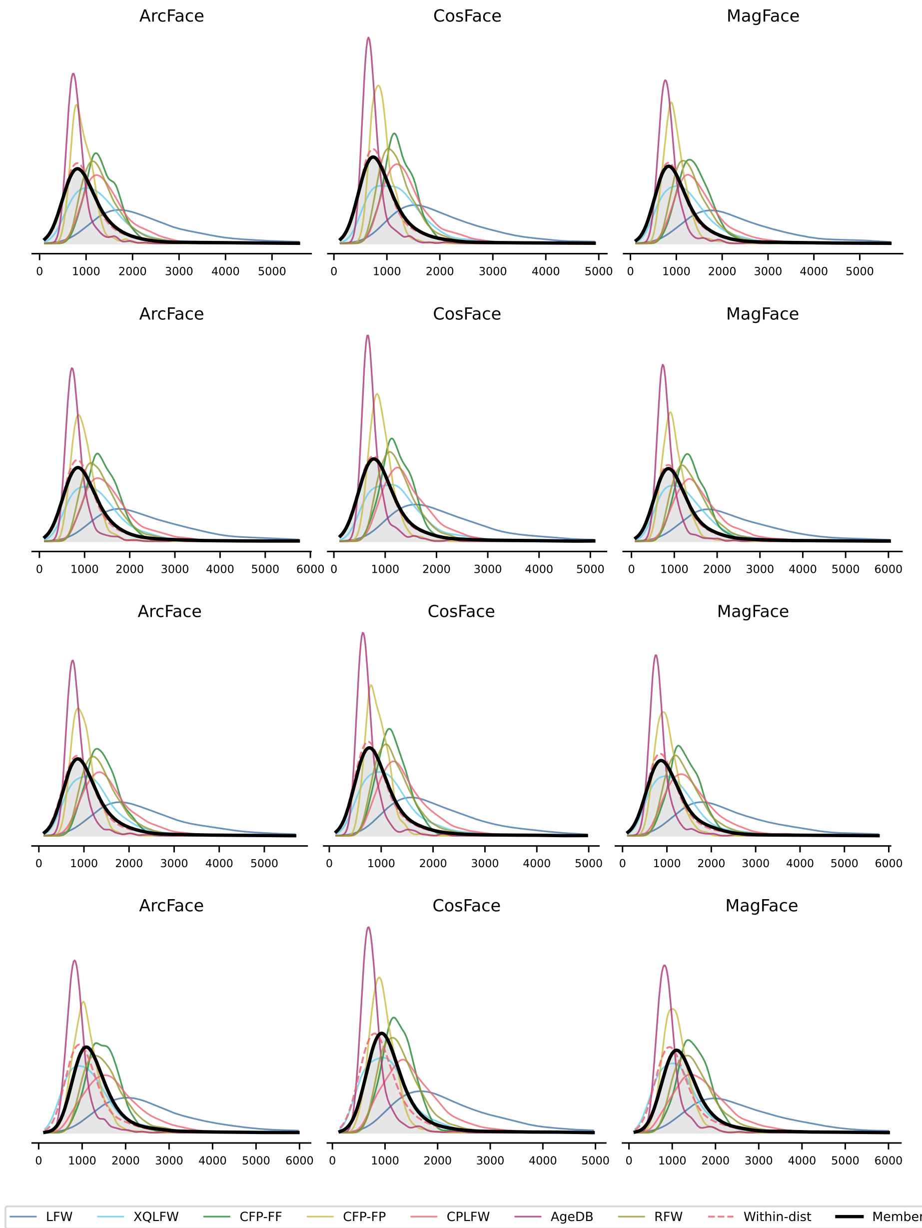


Figure 189. **Vary loss head — IResNet-34 / $n_{\text{ids}} = 100K$.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

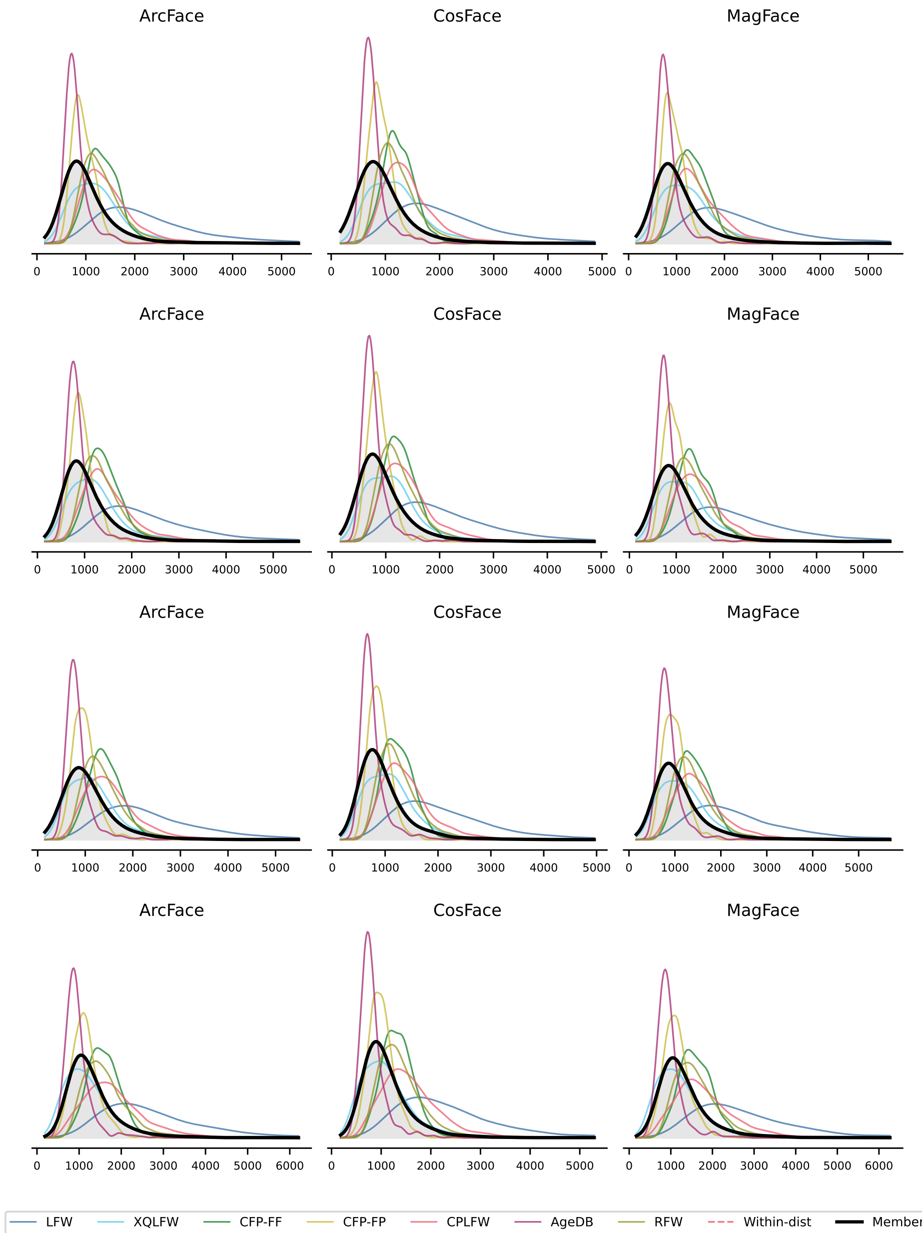


Figure 190. **Vary loss head — IResNet-34 /** $n_{\text{ids}} = All$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

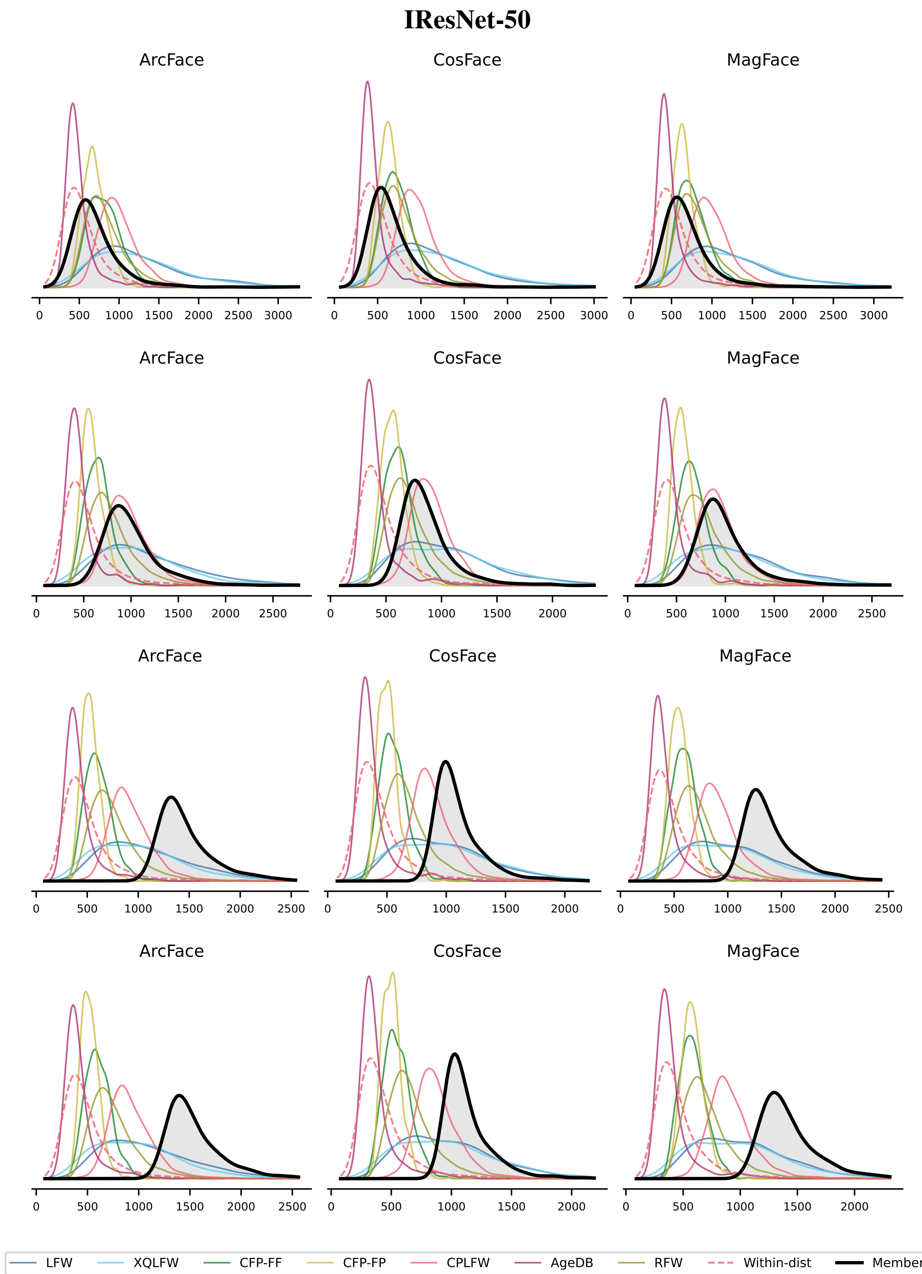


Figure 191. **Vary loss head — IResNet-50 /** $n_{\text{ids}} = 1K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

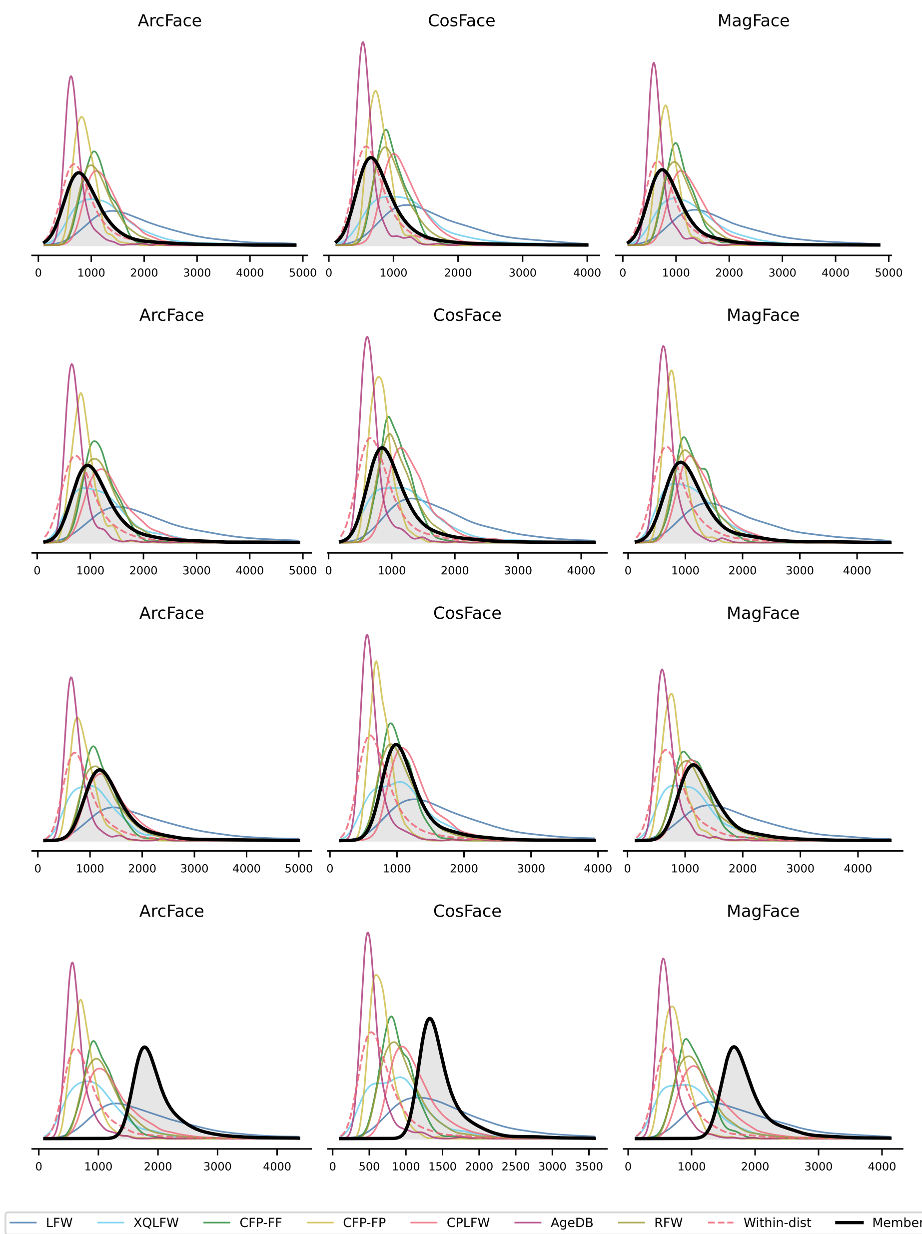


Figure 192. **Vary loss head — IResNet-50 /** $n_{\text{ids}} = 10K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

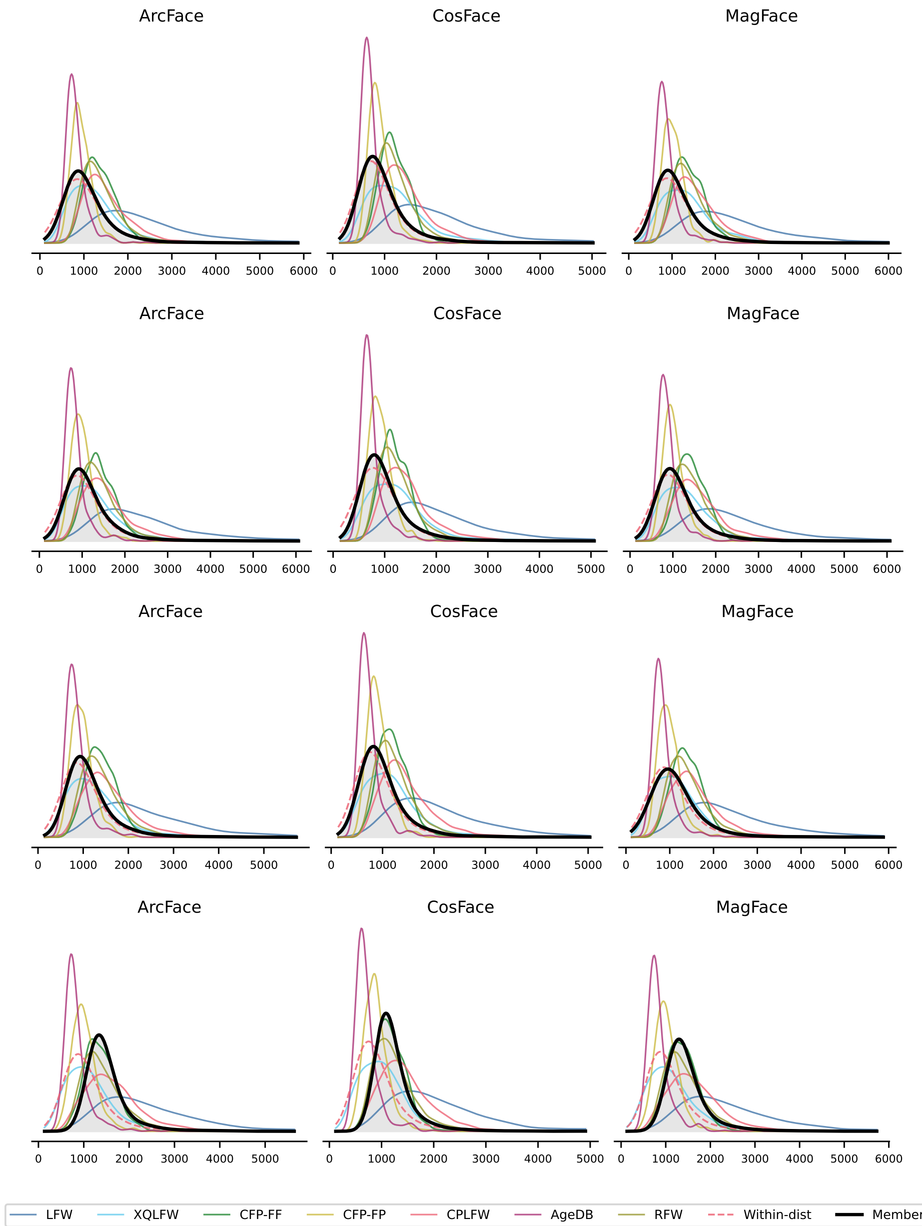


Figure 193. **Vary loss head — IResNet-50 / $n_{\text{ids}} = 50K$.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

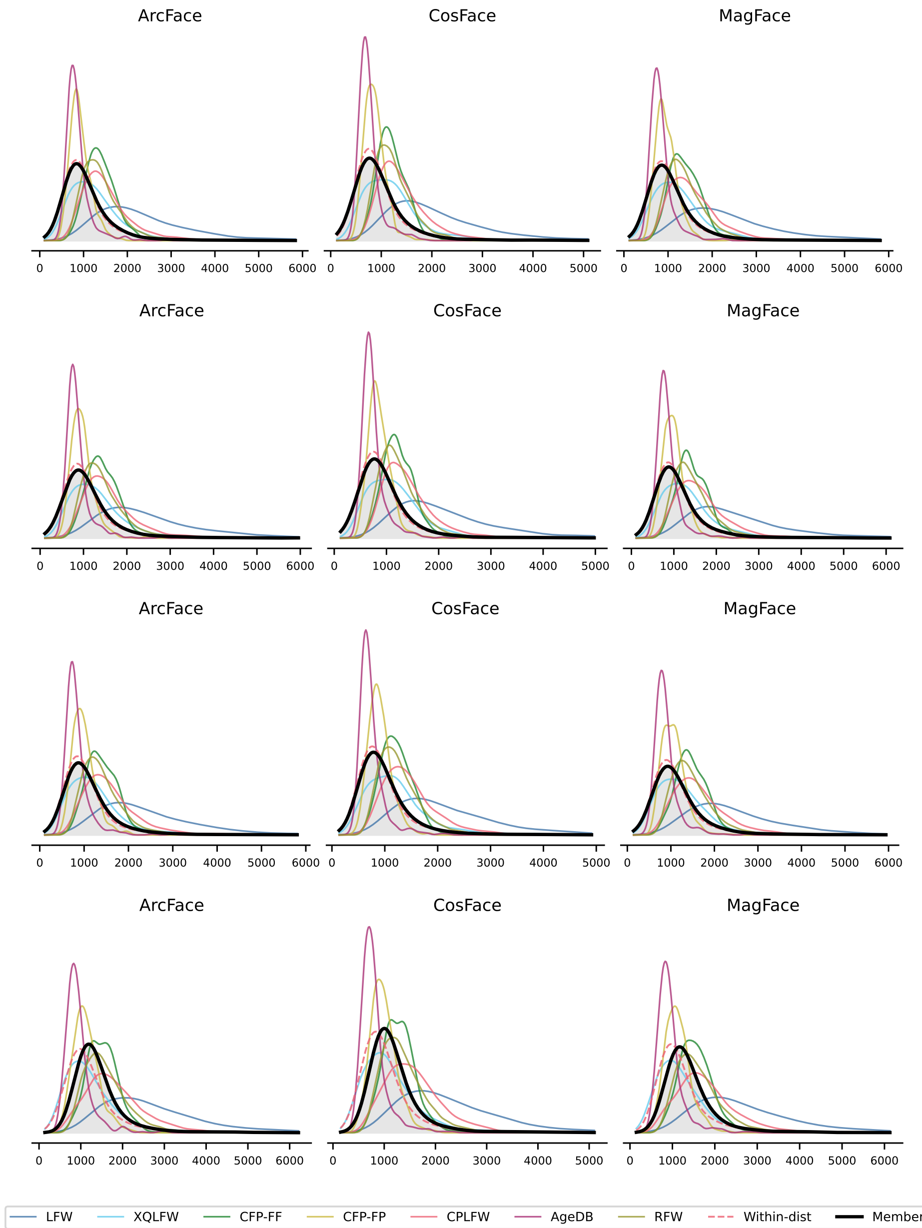


Figure 194. **Vary loss head — IResNet-50 / $n_{\text{ids}} = 100K$.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

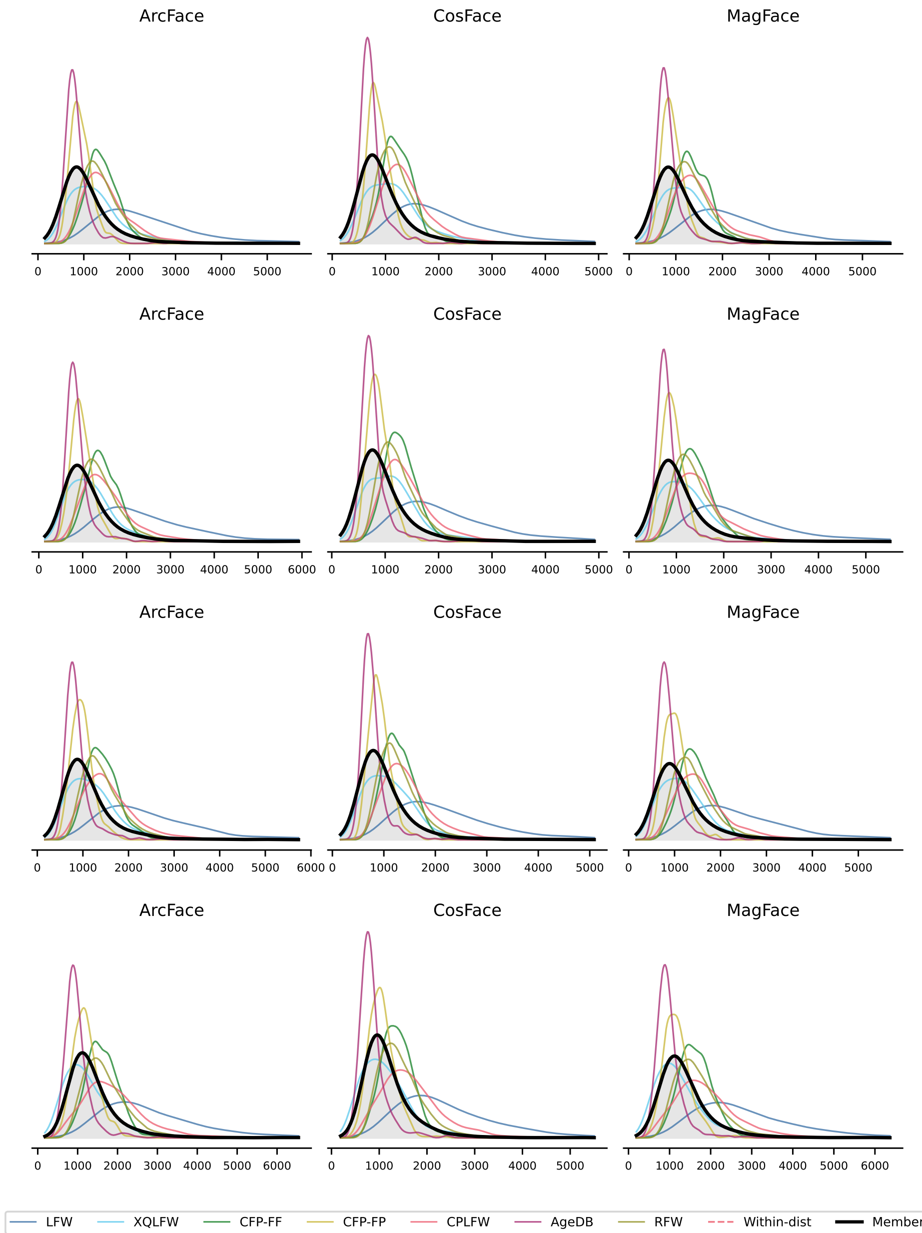


Figure 195. **Vary loss head — IResNet-50 /** $n_{\text{ids}} = All$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

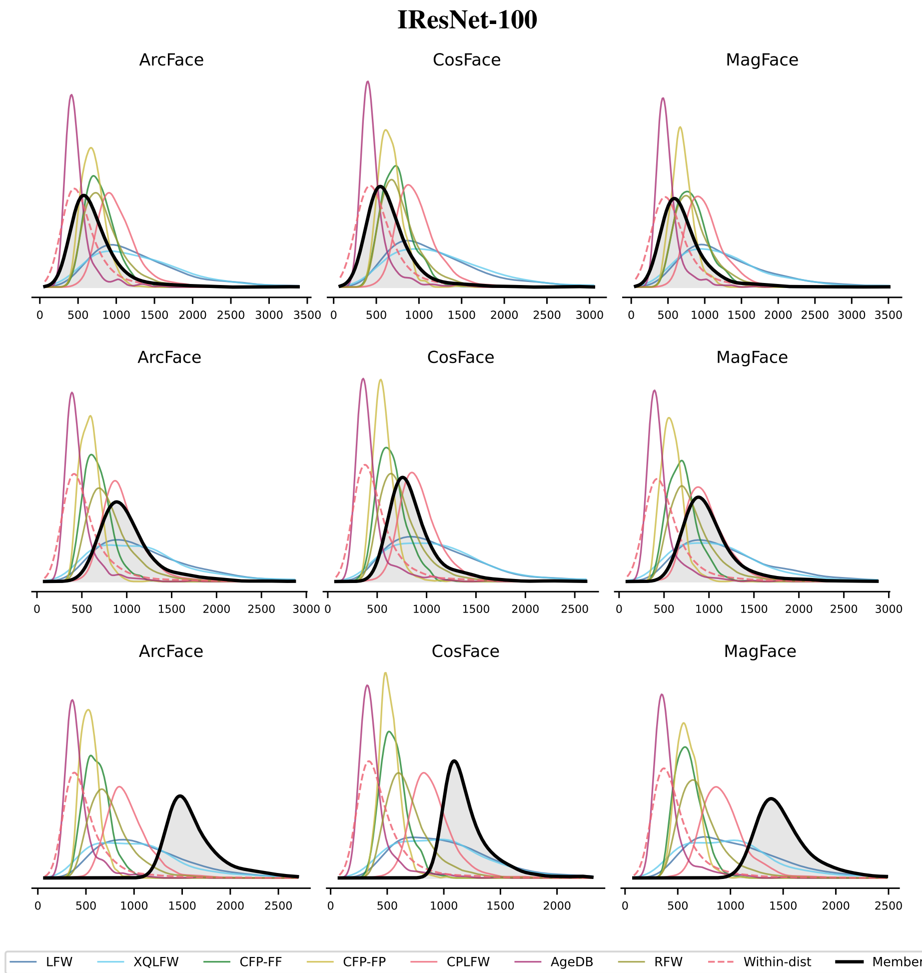


Figure 196. **Vary loss head — IResNet-100 / $n_{\text{ids}} = 1K$.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

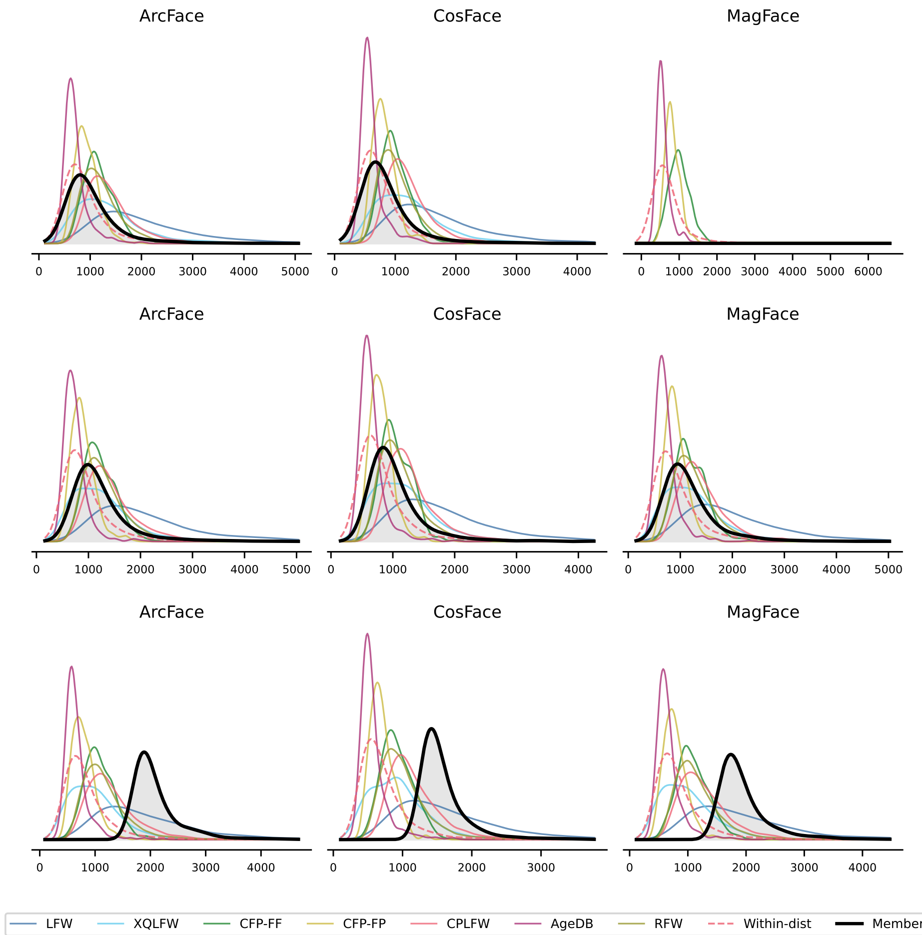


Figure 197. **Vary loss head — IResNet-100 /** $n_{\text{ids}} = 10K$**.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

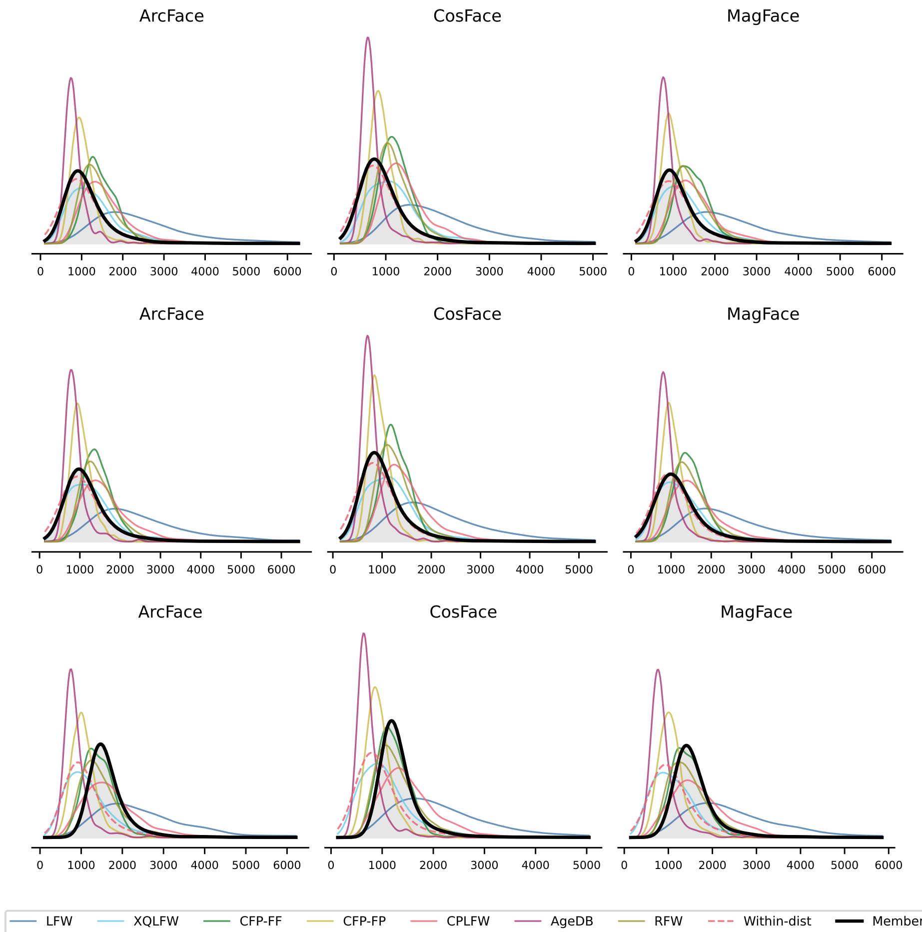


Figure 198. **Vary loss head — IResNet-100 /** $n_{\text{ids}} = 50K$**.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

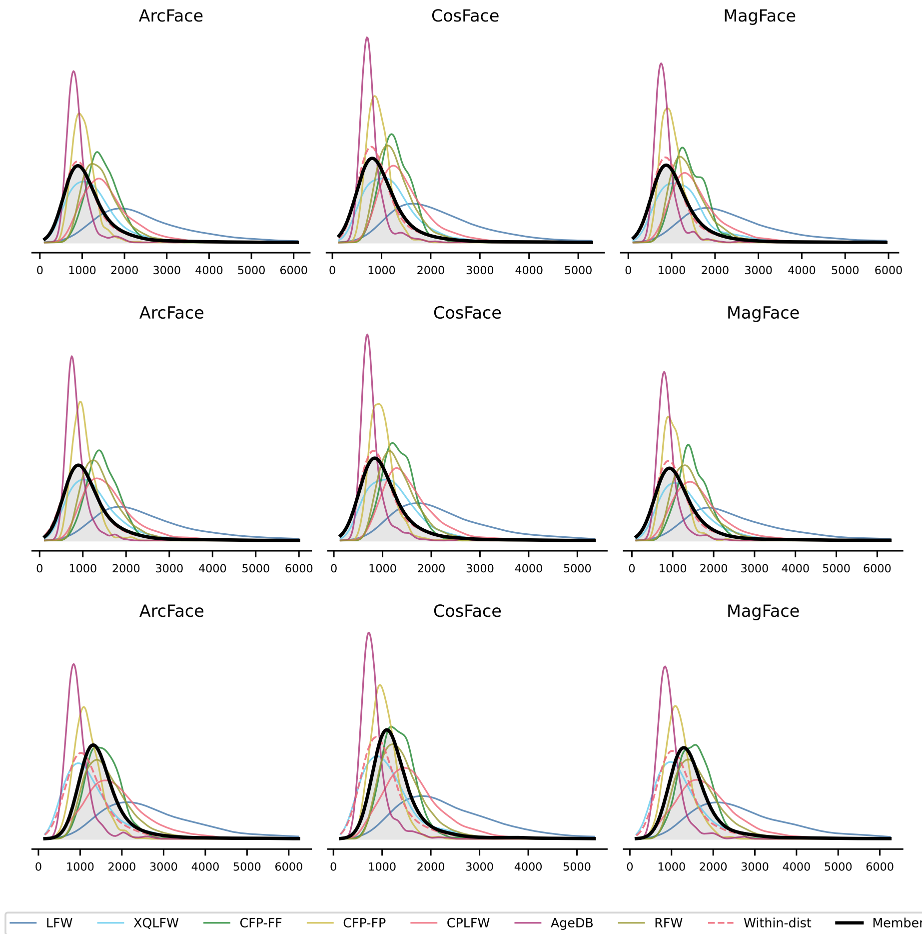


Figure 199. **Vary loss head — IResNet-100 /** $n_{\textbf{ids}} = 100K$**.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

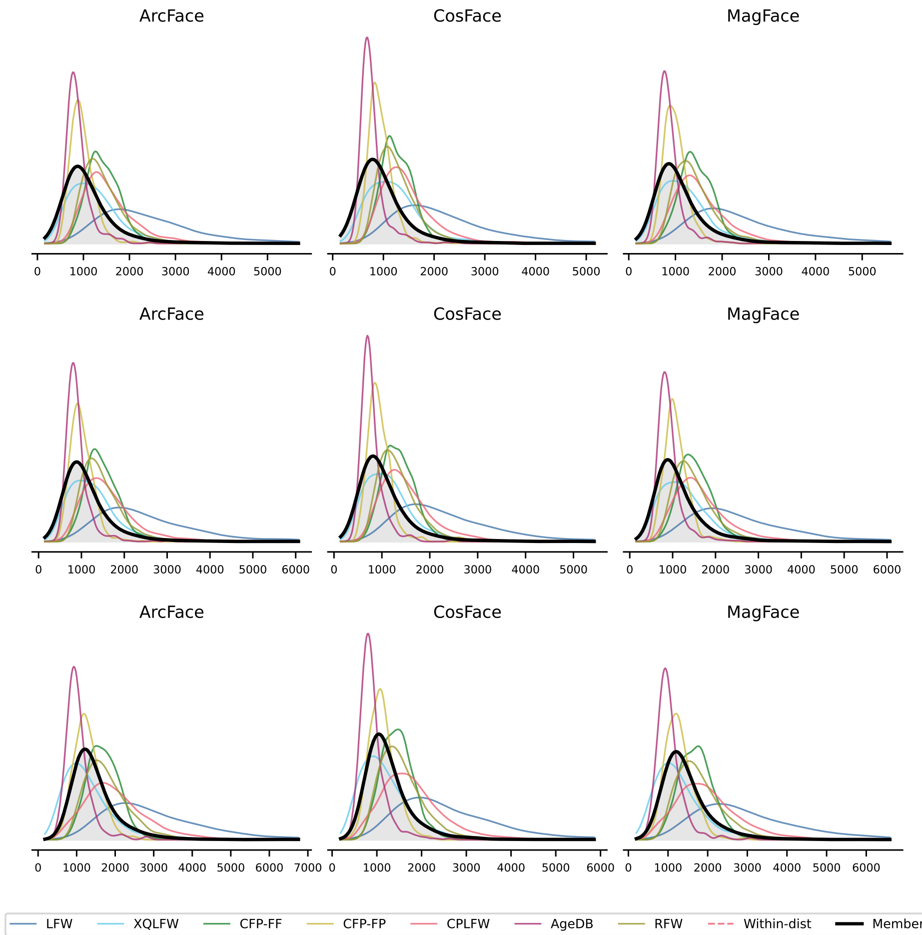


Figure 200. **Vary loss head — IResNet-100 / $n_{\textbf{ids}} = All$.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

5.4. Vary backbone

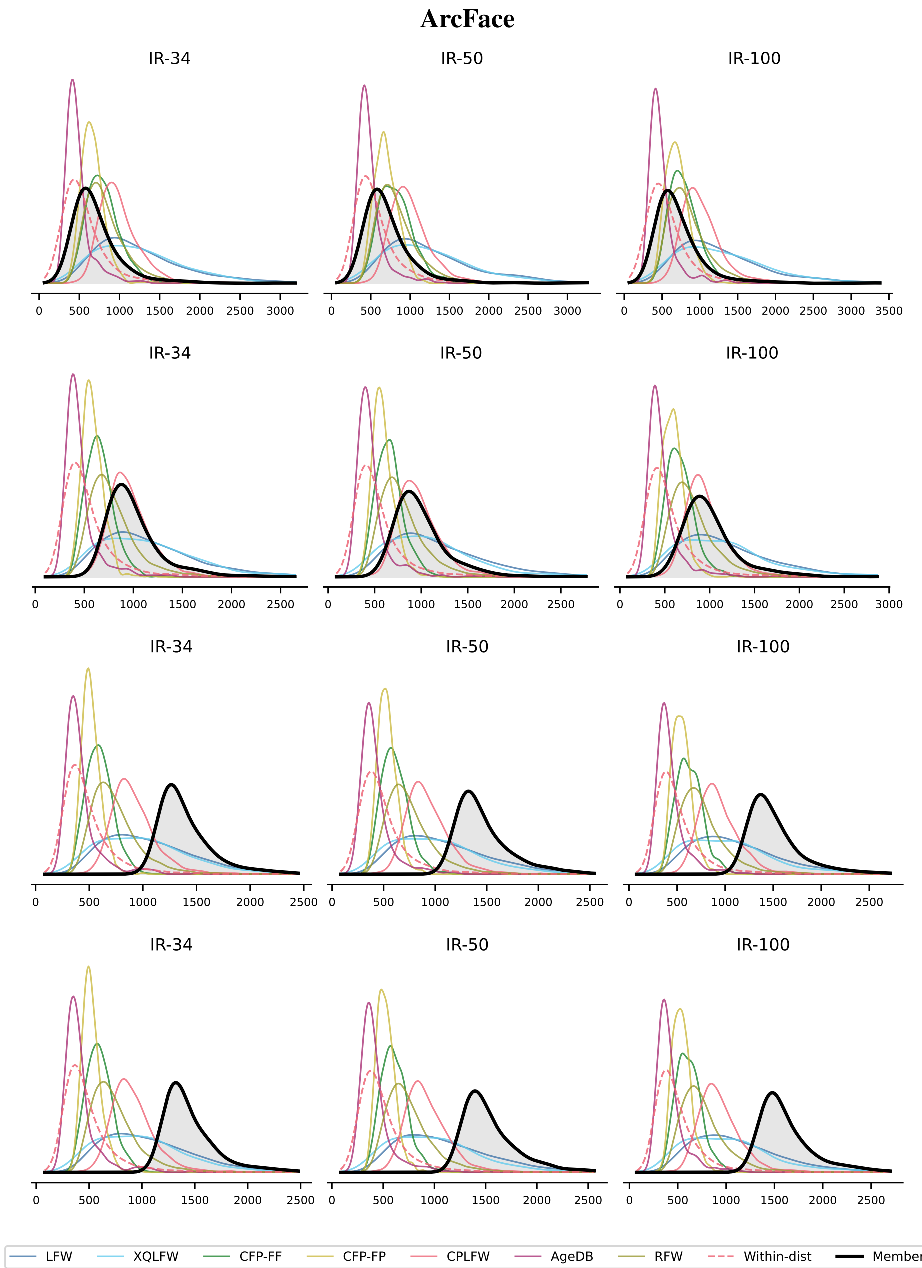


Figure 201. **Vary backbone — ArcFace /** $n_{\text{ids}} = 1K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

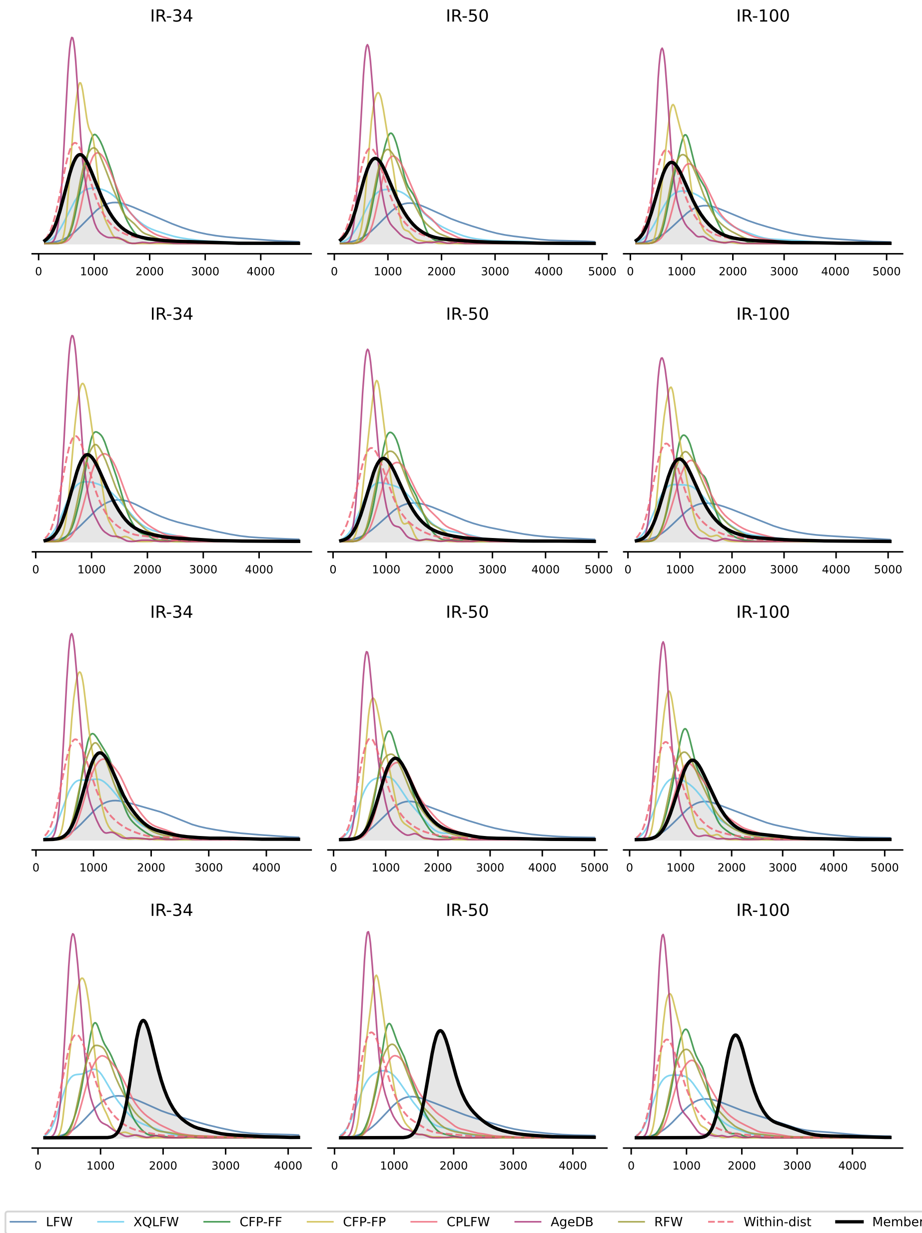


Figure 202. **Vary backbone — ArcFace /** $n_{\textbf{ids}} = 10K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

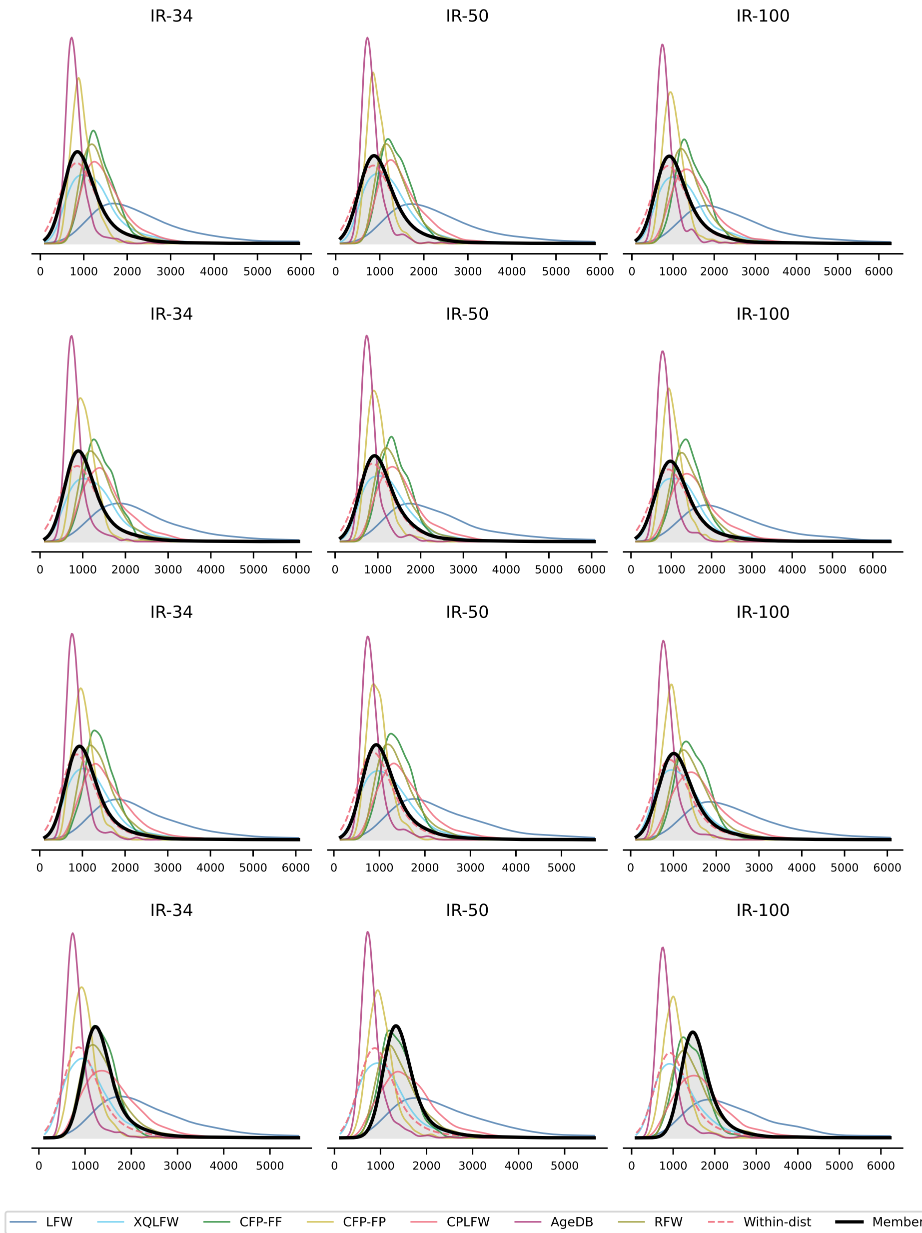


Figure 203. **Vary backbone — ArcFace /** $n_{\mathbf{ids}} = 50K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

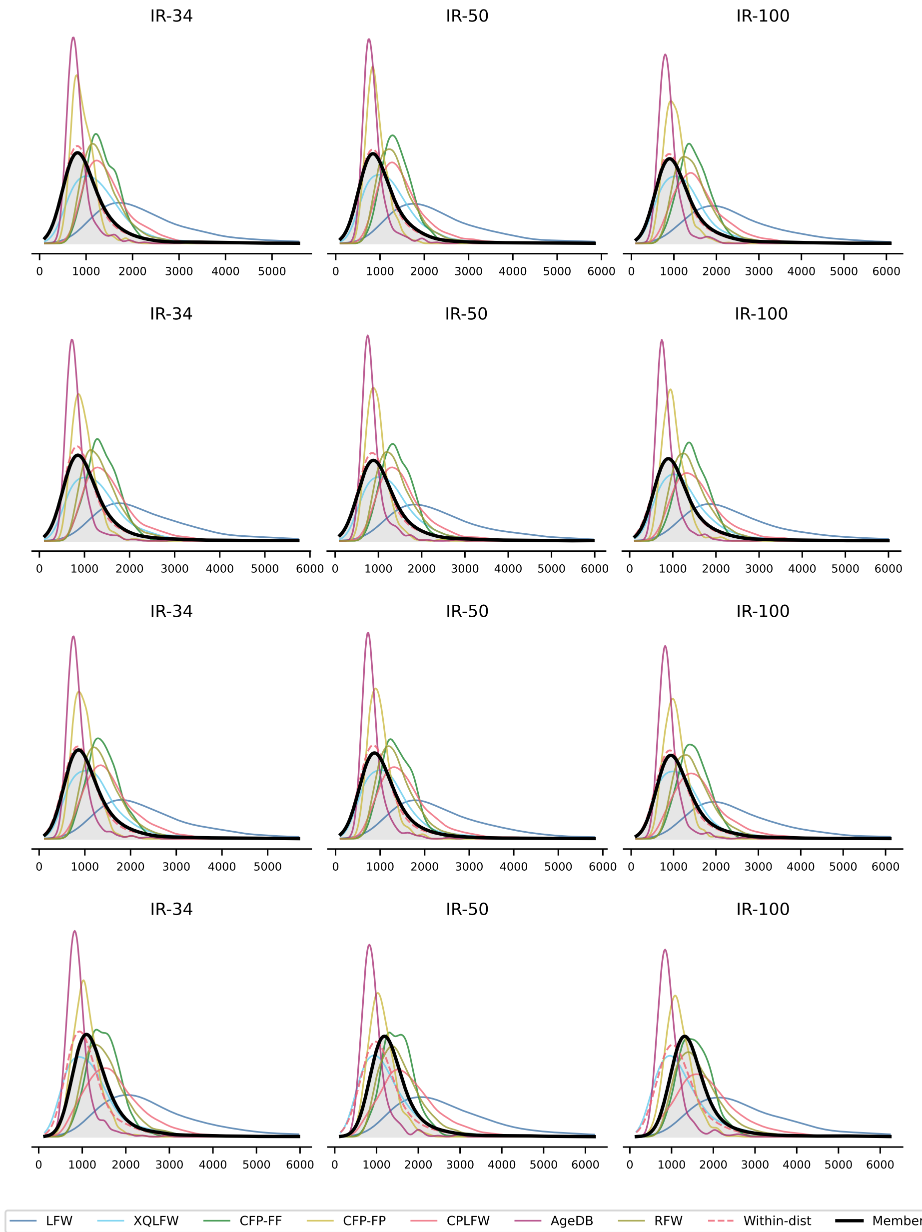


Figure 204. **Vary backbone — ArcFace /** $n_{\text{ids}} = 100K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

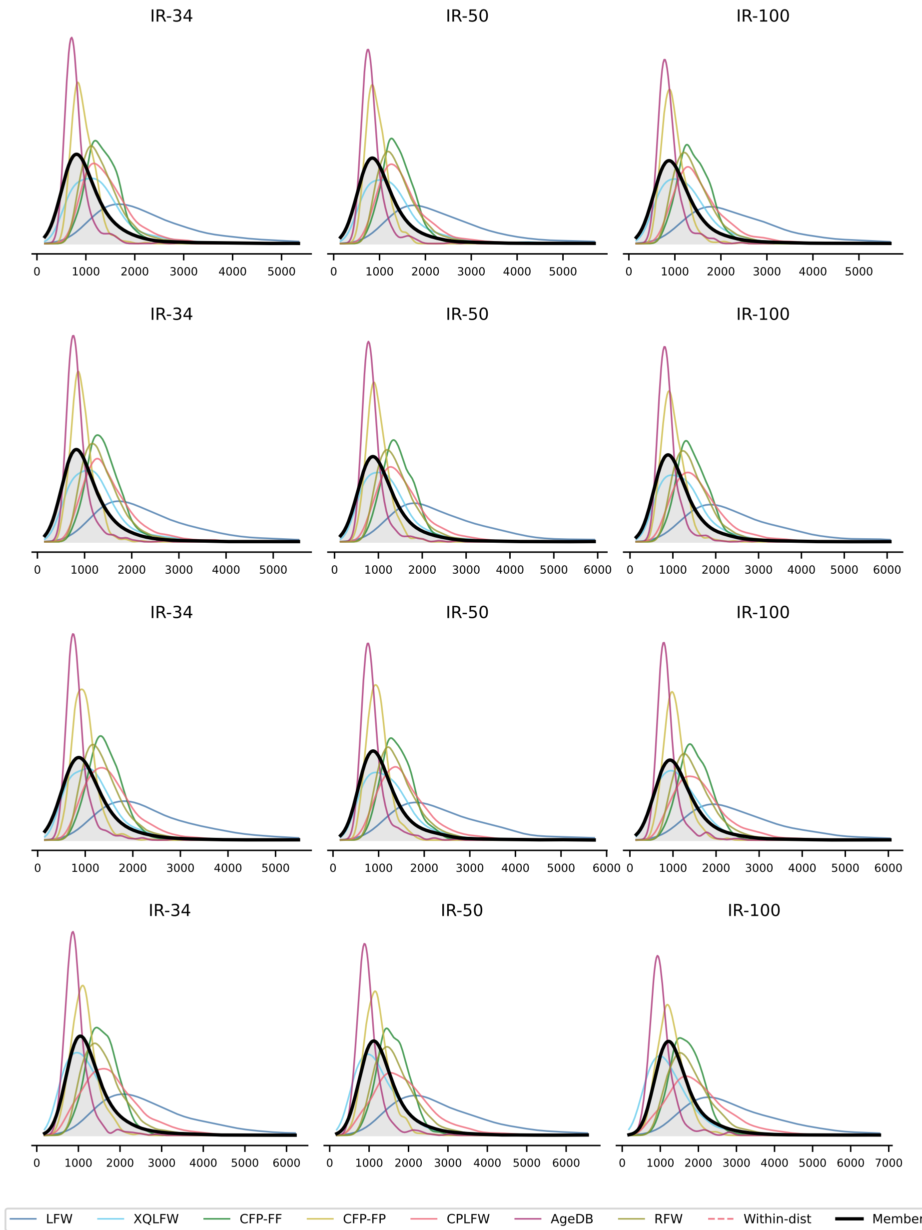


Figure 205. **Vary backbone — ArcFace /** $n_{\text{ids}} = All$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

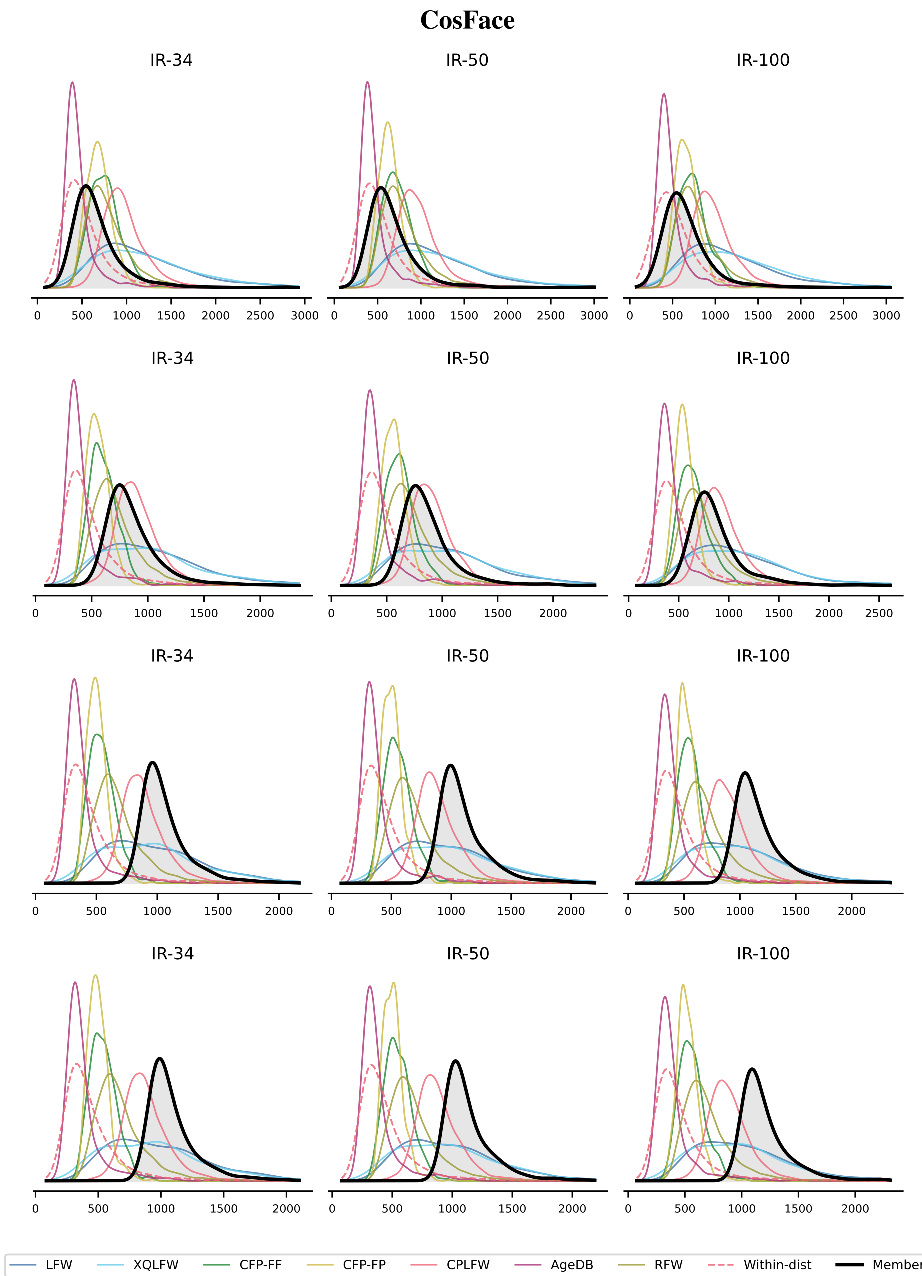


Figure 206. **Vary backbone — CosFace /** $n_{\text{ids}} = 1K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

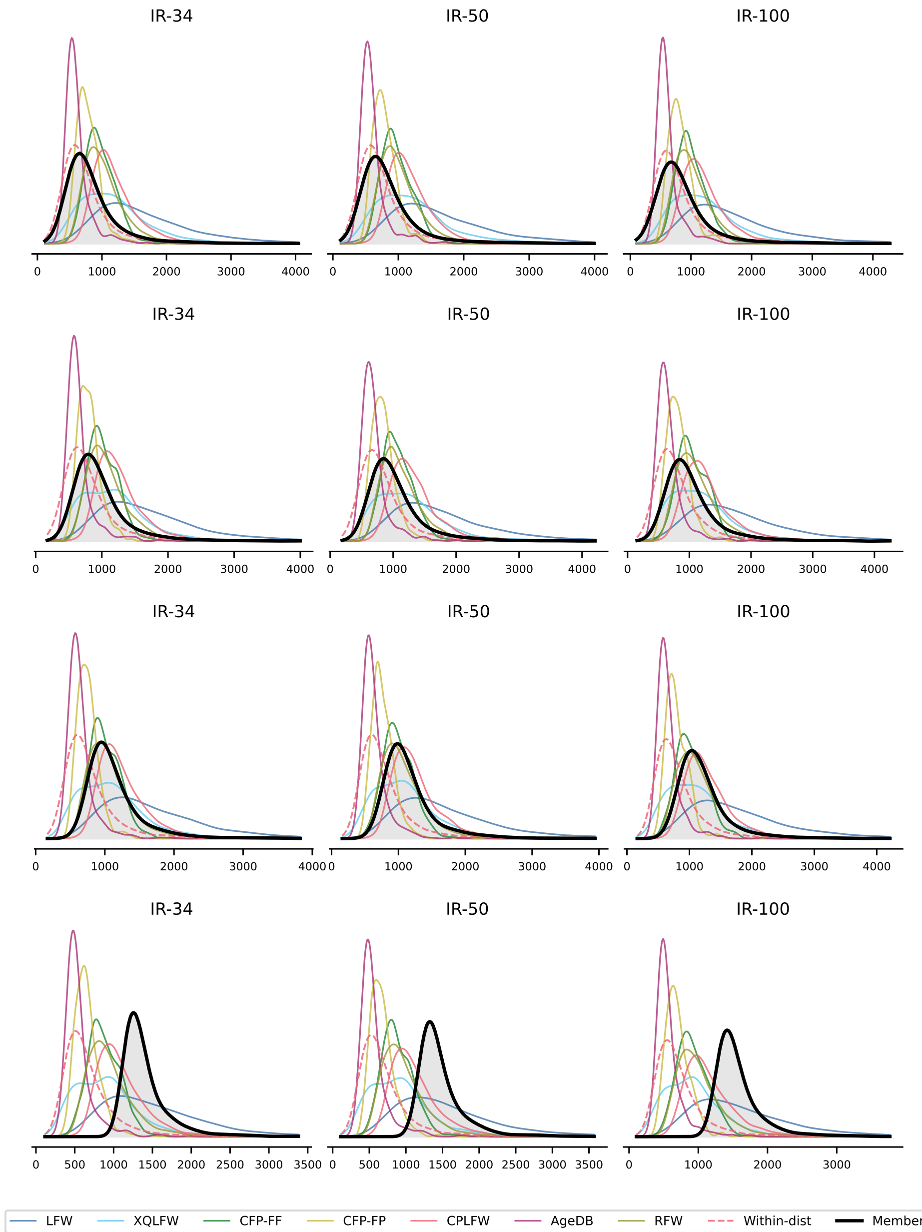


Figure 207. **Vary backbone — CosFace /** $n_{\mathbf{ids}} = 10K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

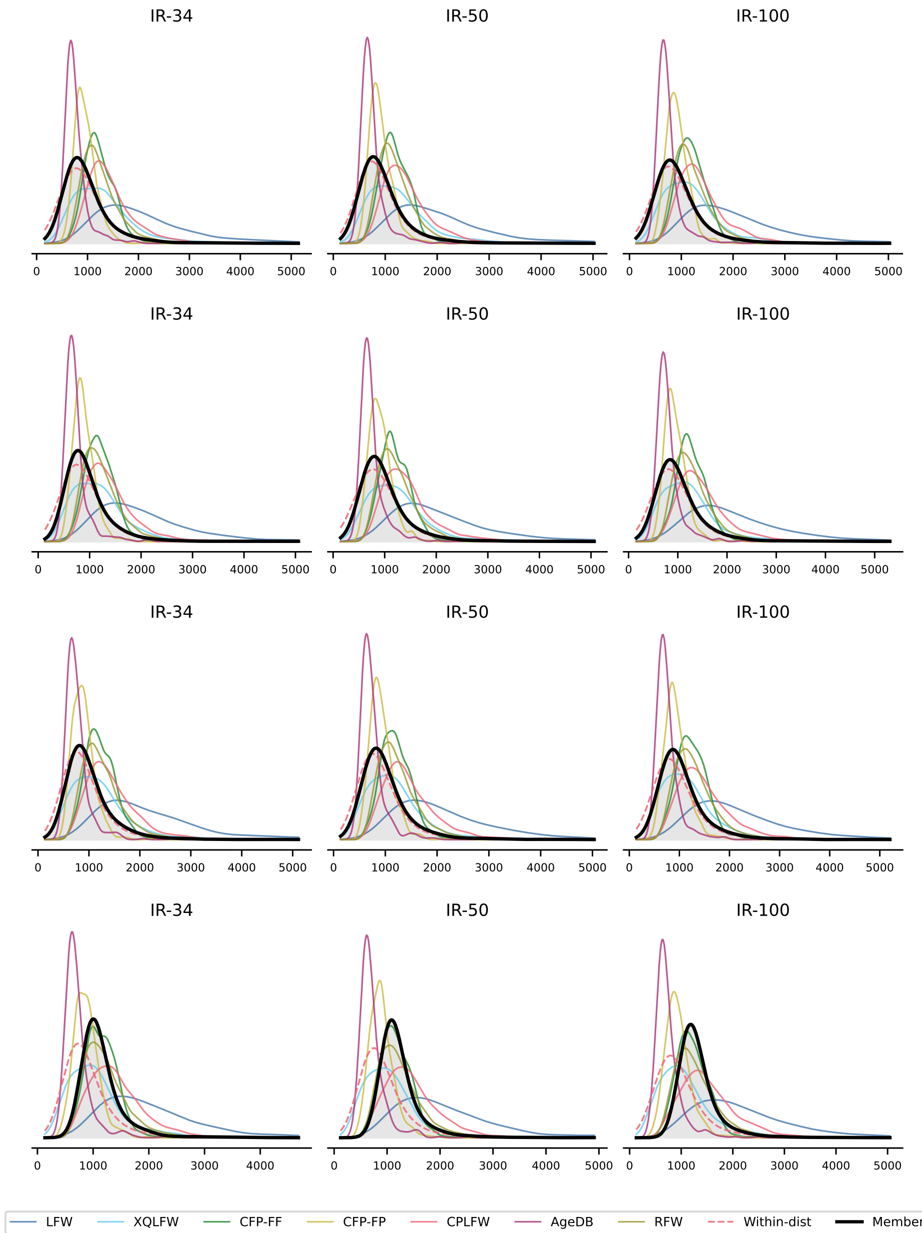


Figure 208. **Vary backbone — CosFace /** $n_{\textbf{ids}} = 50K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

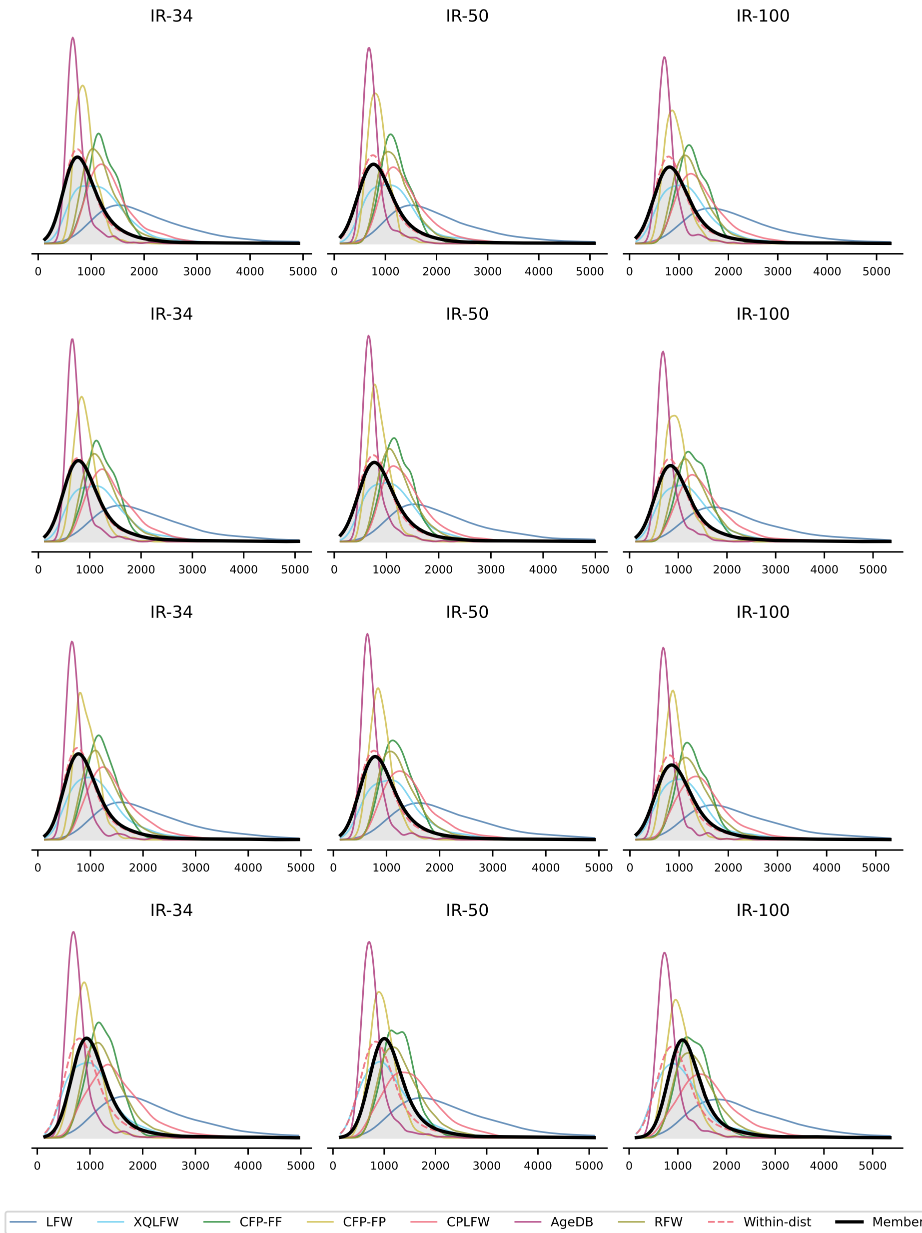


Figure 209. **Vary backbone — CosFace /** $n_{\mathbf{ids}} = 100K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

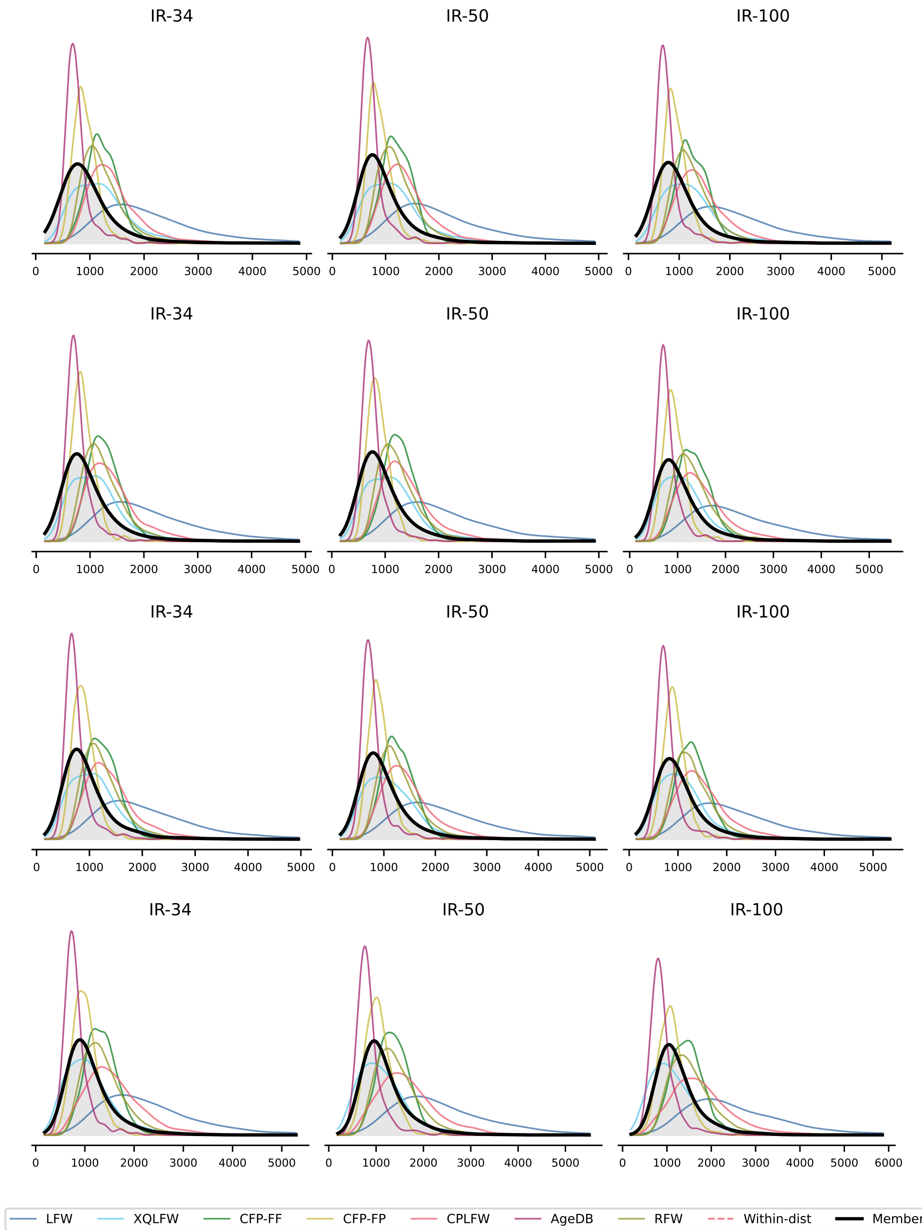


Figure 210. **Vary backbone — CosFace / $n_{\text{ids}} = All$.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

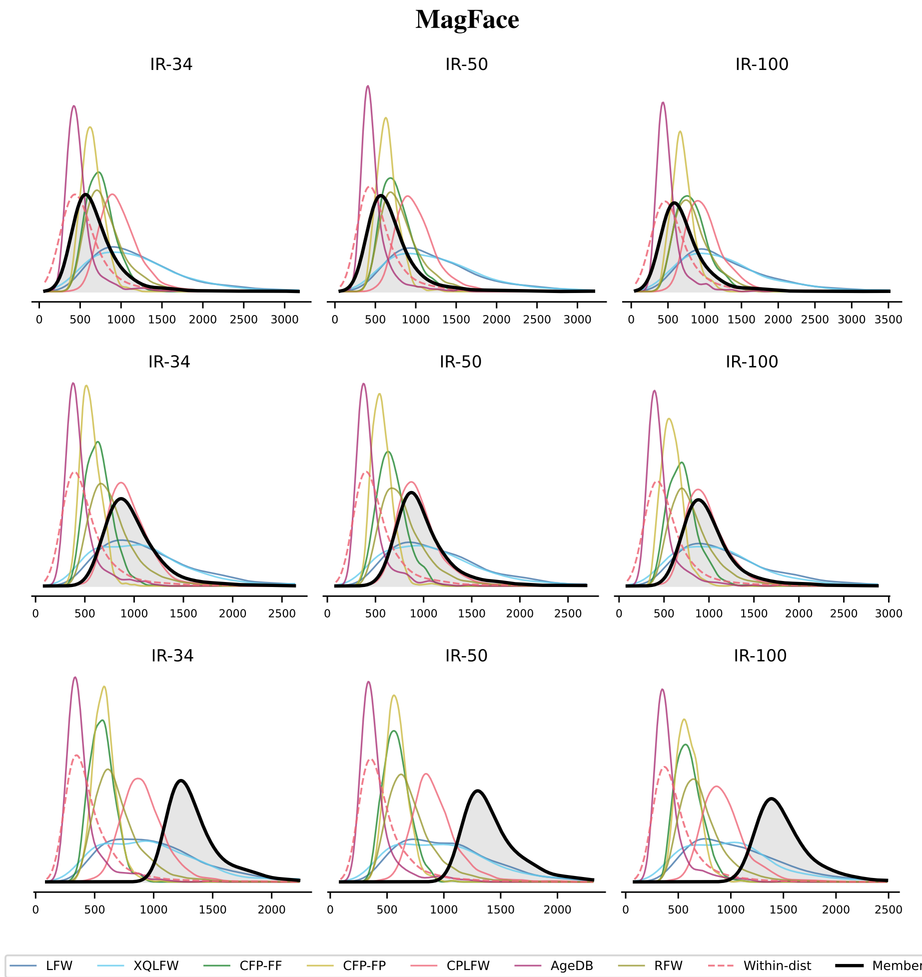


Figure 211. **Vary backbone — MagFace / $n_{\text{ids}} = 1K$.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

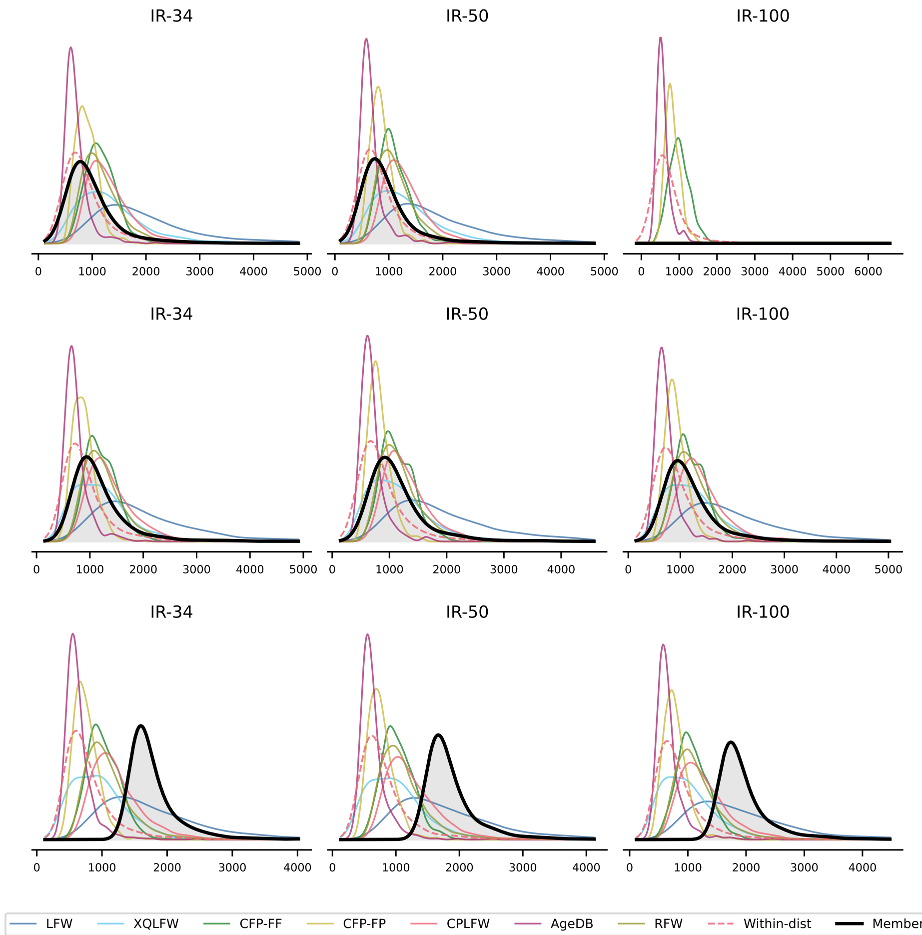


Figure 212. **Vary backbone — MagFace /** $n_{\text{ids}} = 10K$**.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

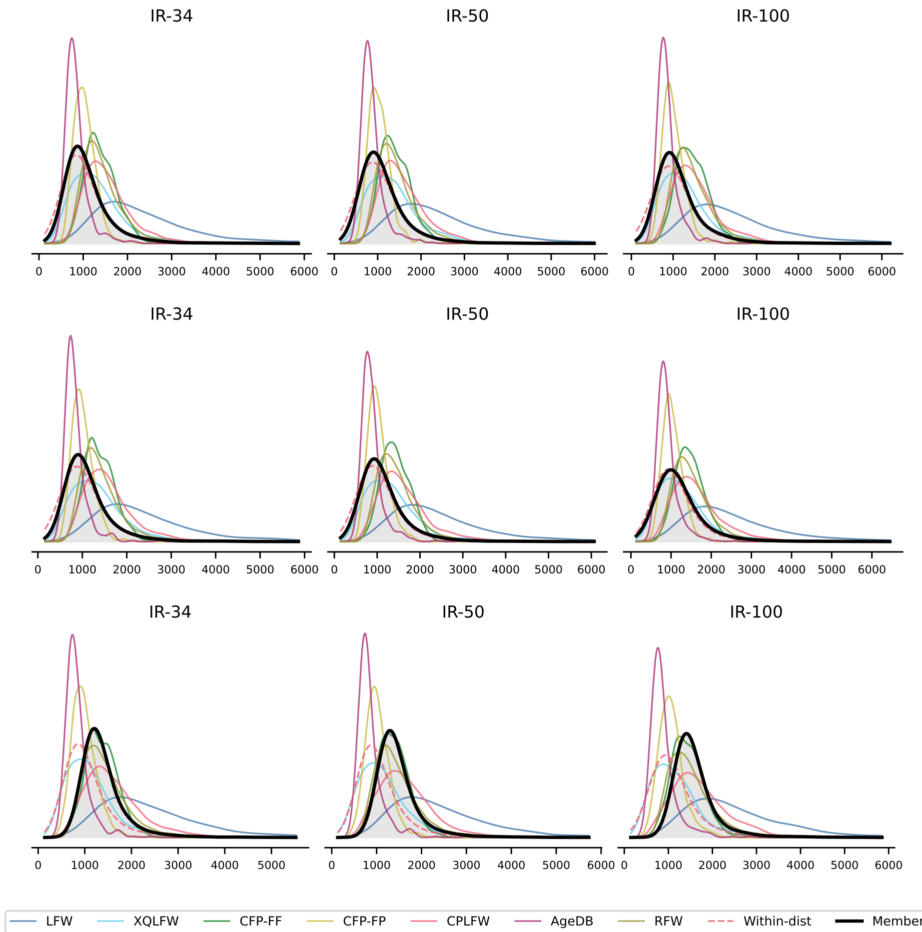


Figure 213. **Vary backbone — MagFace /** $n_{\text{ids}} = 50K$**.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

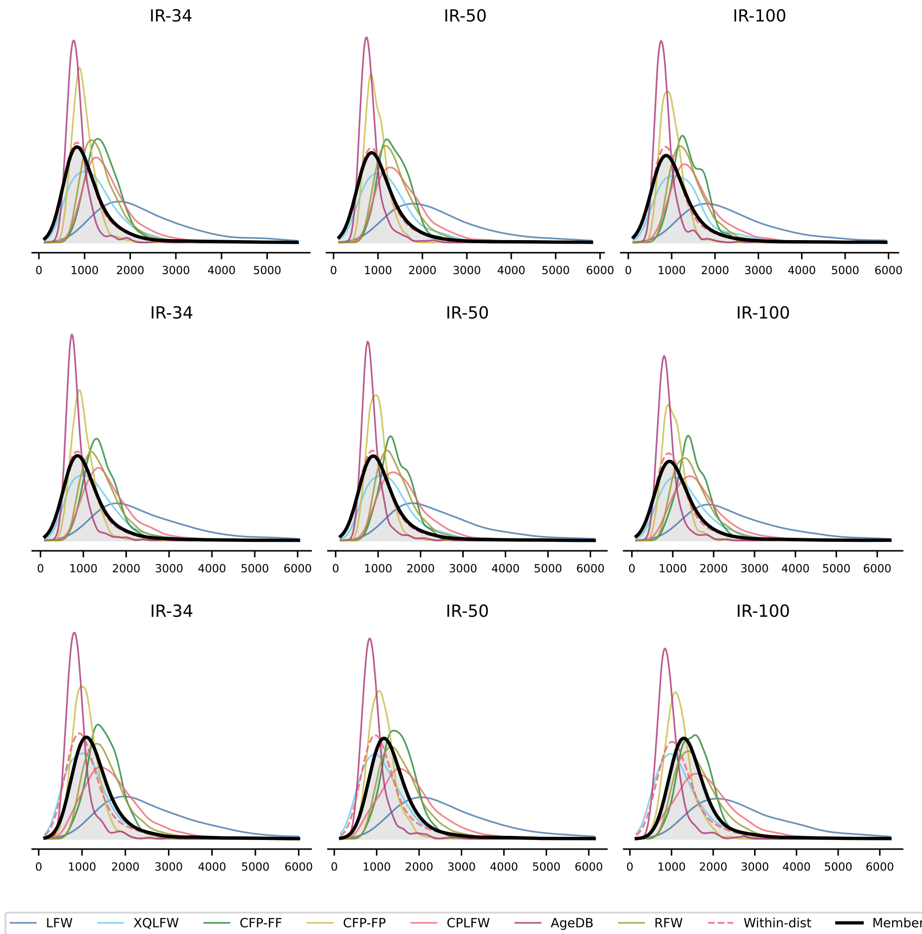


Figure 214. **Vary backbone — MagFace / $n_{\text{ids}} = 100K$.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

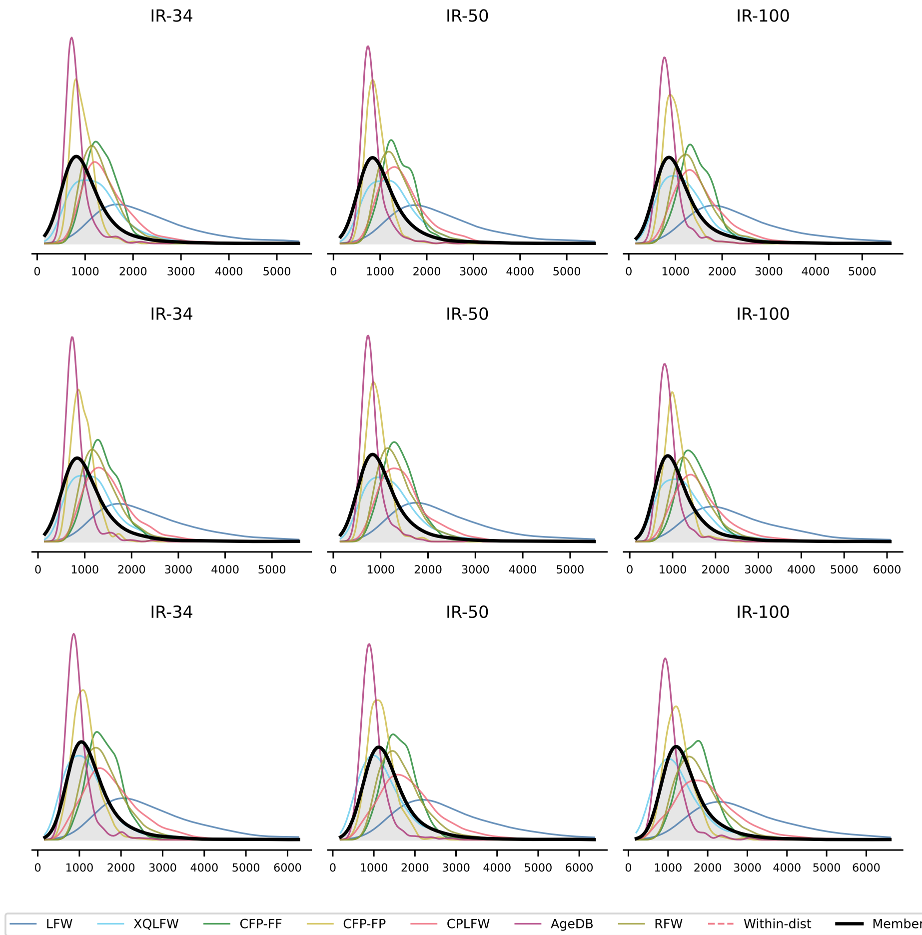


Figure 215. **Vary backbone — MagFace / $n_{\mathbf{ids}} = All$.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

6. Penalized logit distributions

6.1. Vary training duration

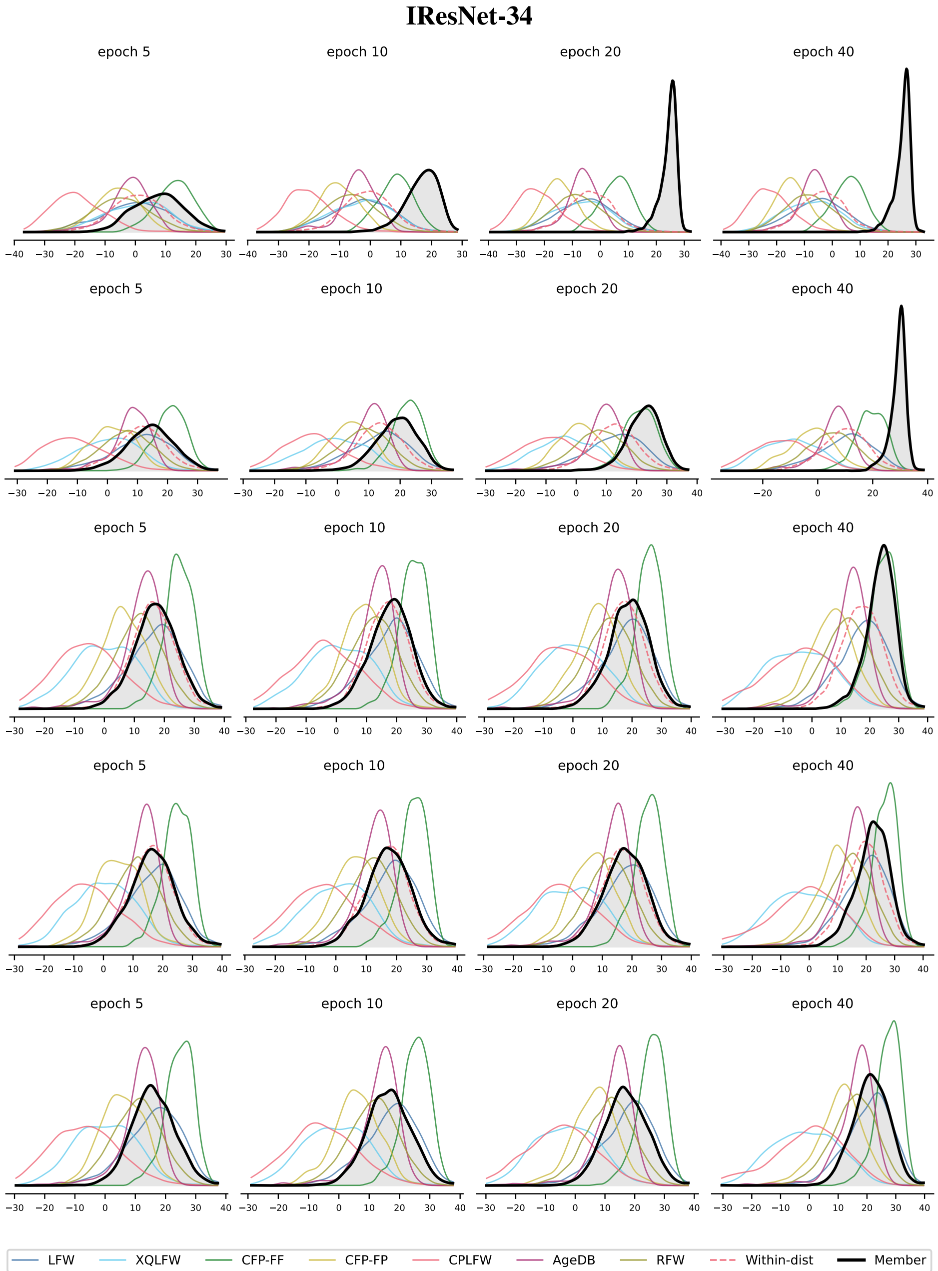


Figure 216. **Vary training duration — IResNet-34 / ArcFace.** Rows: $n_{\text{ids}} \in \{1\text{K},\ 10\text{K},\ 50\text{K},\ 100\text{K},\ \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

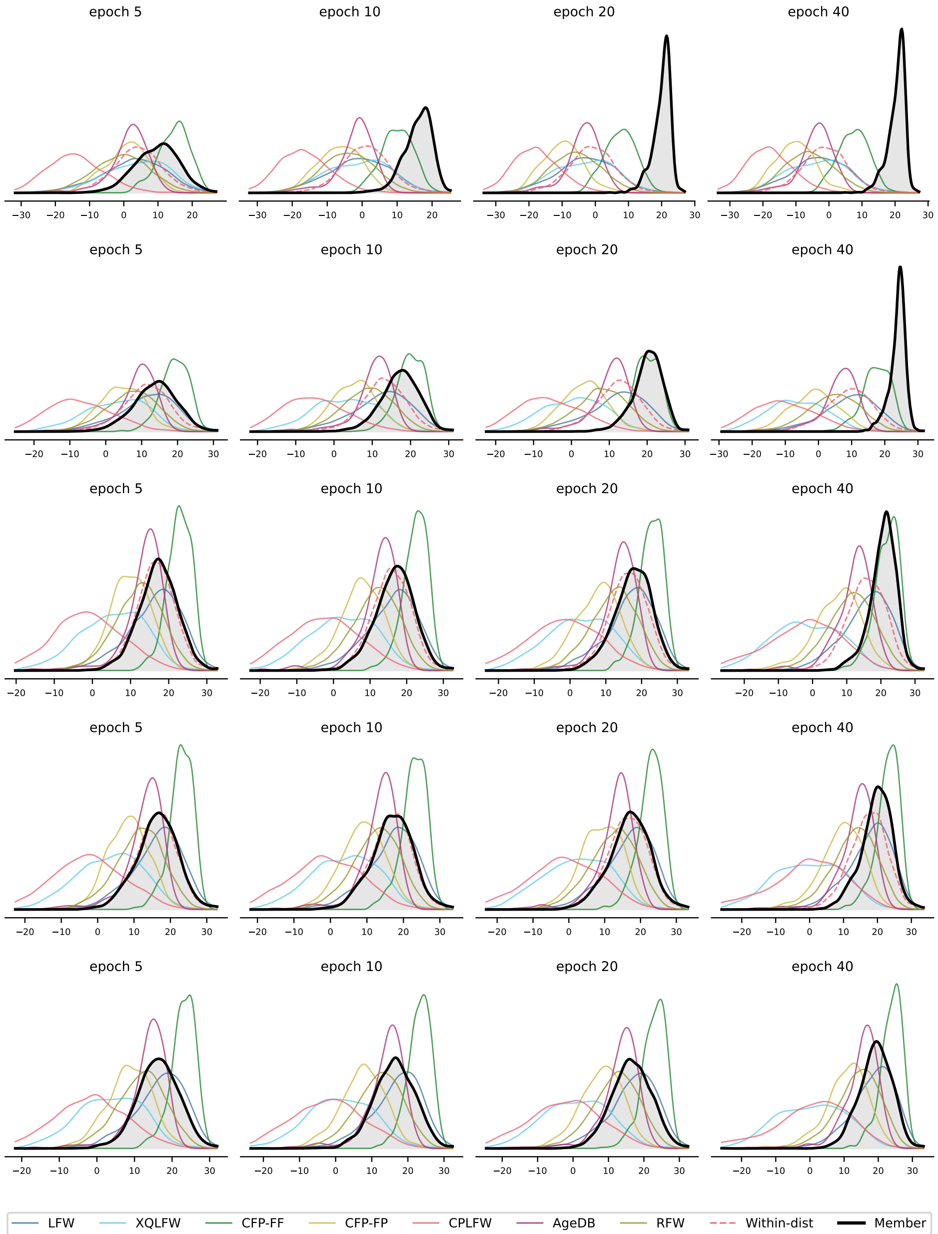


Figure 217. **Vary training duration — IResNet-34 / CosFace.** Rows: $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

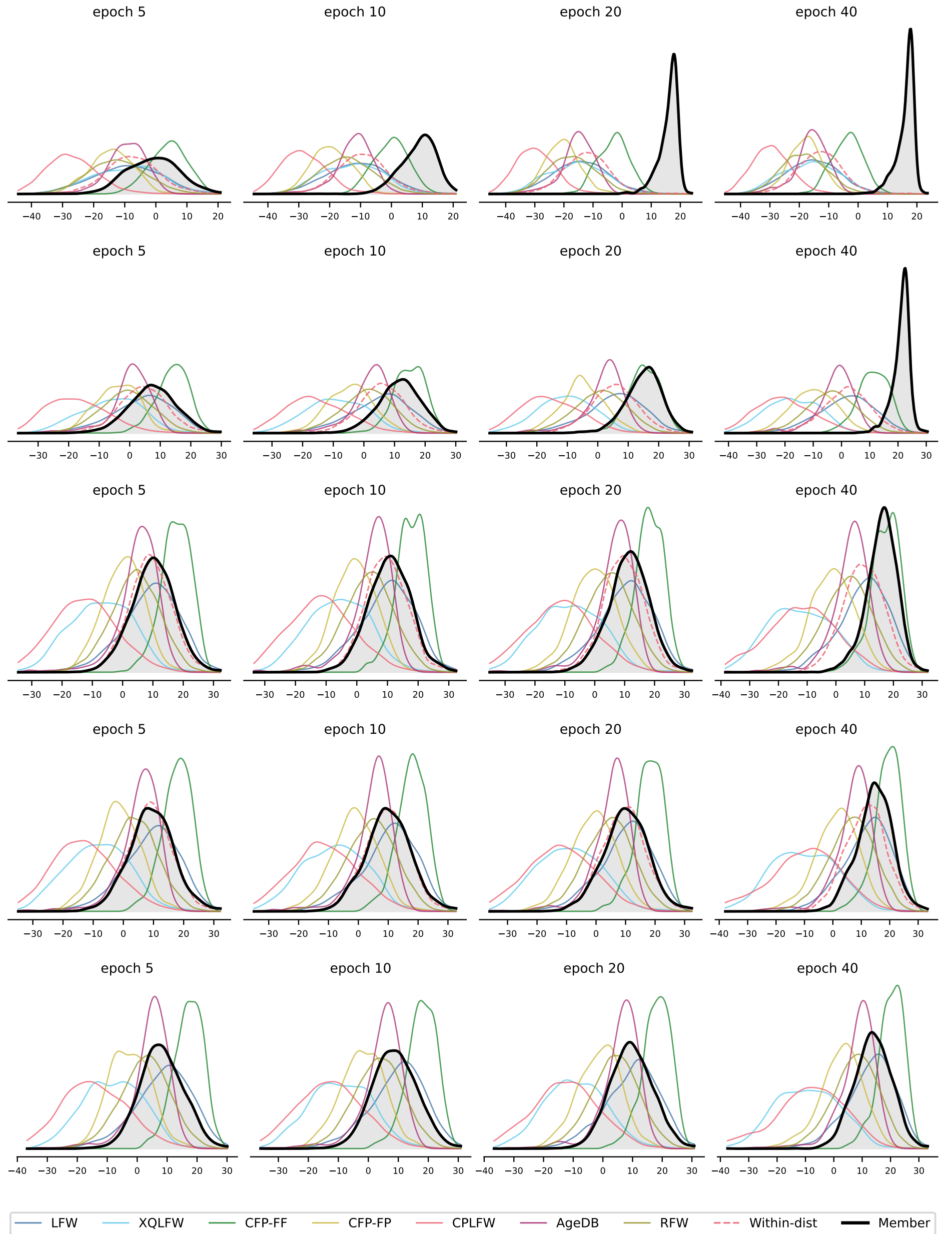


Figure 218. **Vary training duration — IResNet-34 / MagFace.** Rows: $n_{\text{ids}} \in \{1\text{K},\ 10\text{K},\ 50\text{K},\ 100\text{K},\ \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

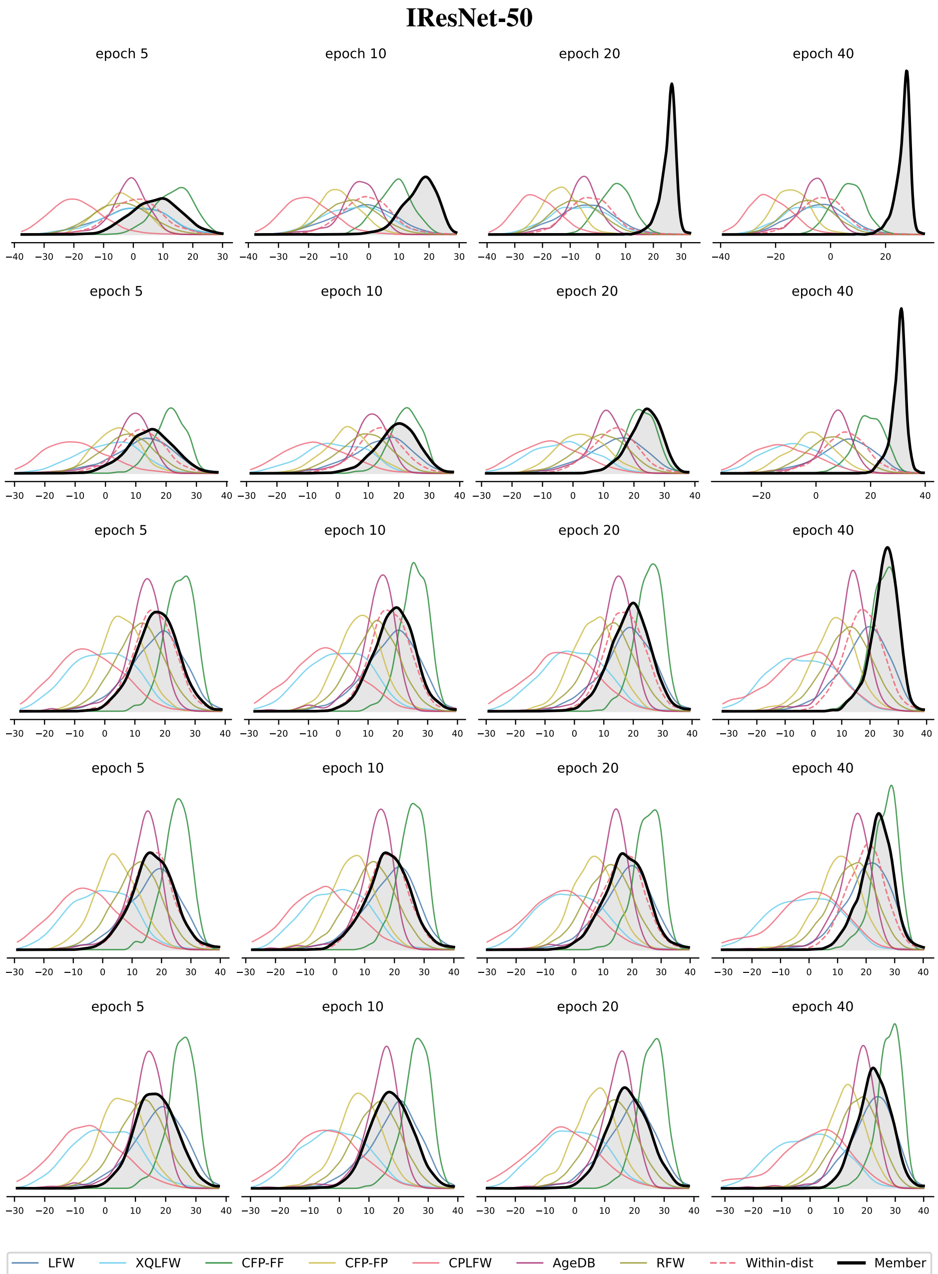


Figure 219. **Vary training duration — IResNet-50 / ArcFace.** Rows: $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

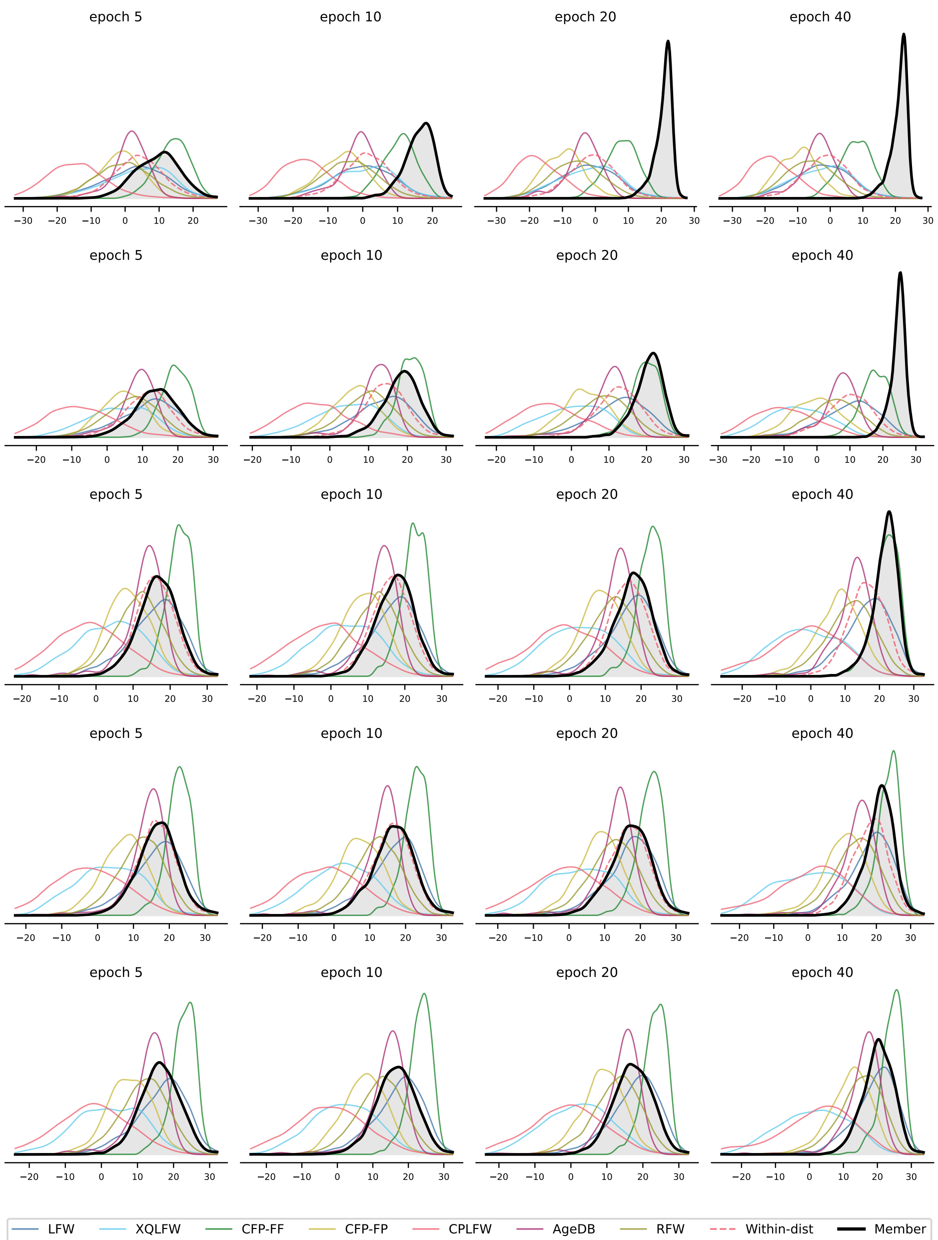


Figure 220. **Vary training duration — IResNet-50 / CosFace.** Rows: $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

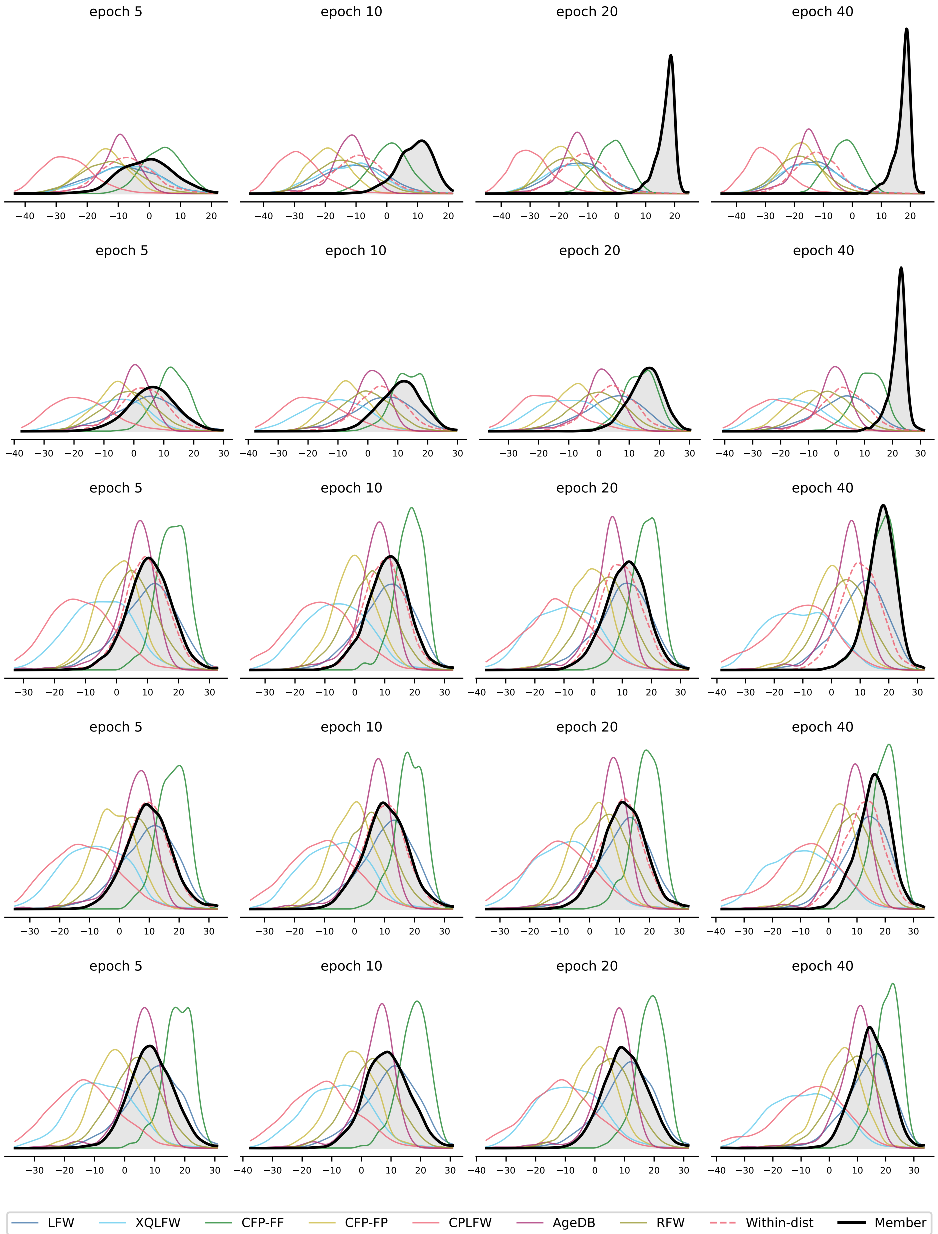


Figure 221. **Vary training duration — IResNet-50 / MagFace.** Rows: $n_{\text{ids}} \in \{1\text{K},\ 10\text{K},\ 50\text{K},\ 100\text{K},\ \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

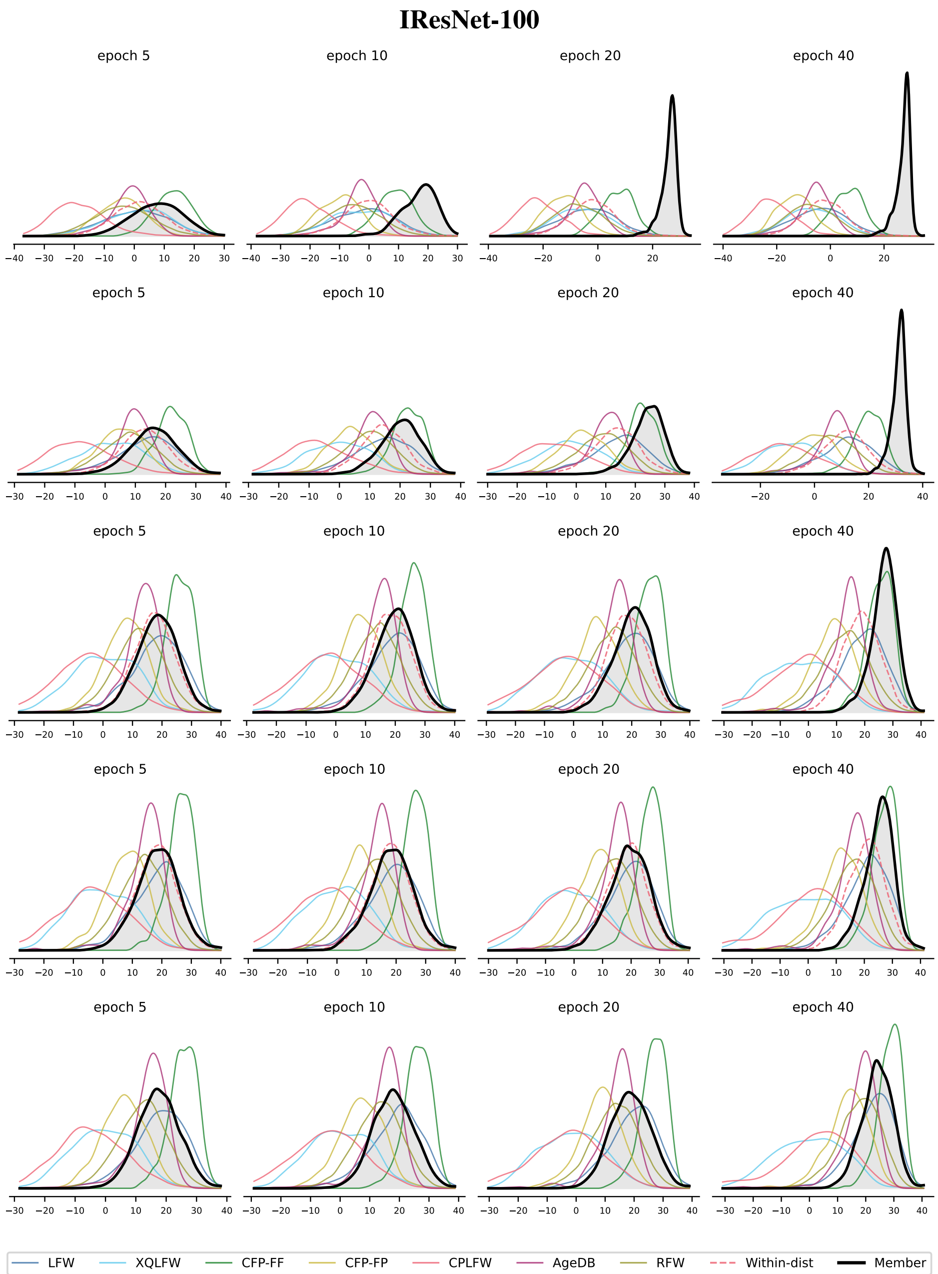


Figure 222. **Vary training duration — IResNet-100 / ArcFace.** Rows: $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

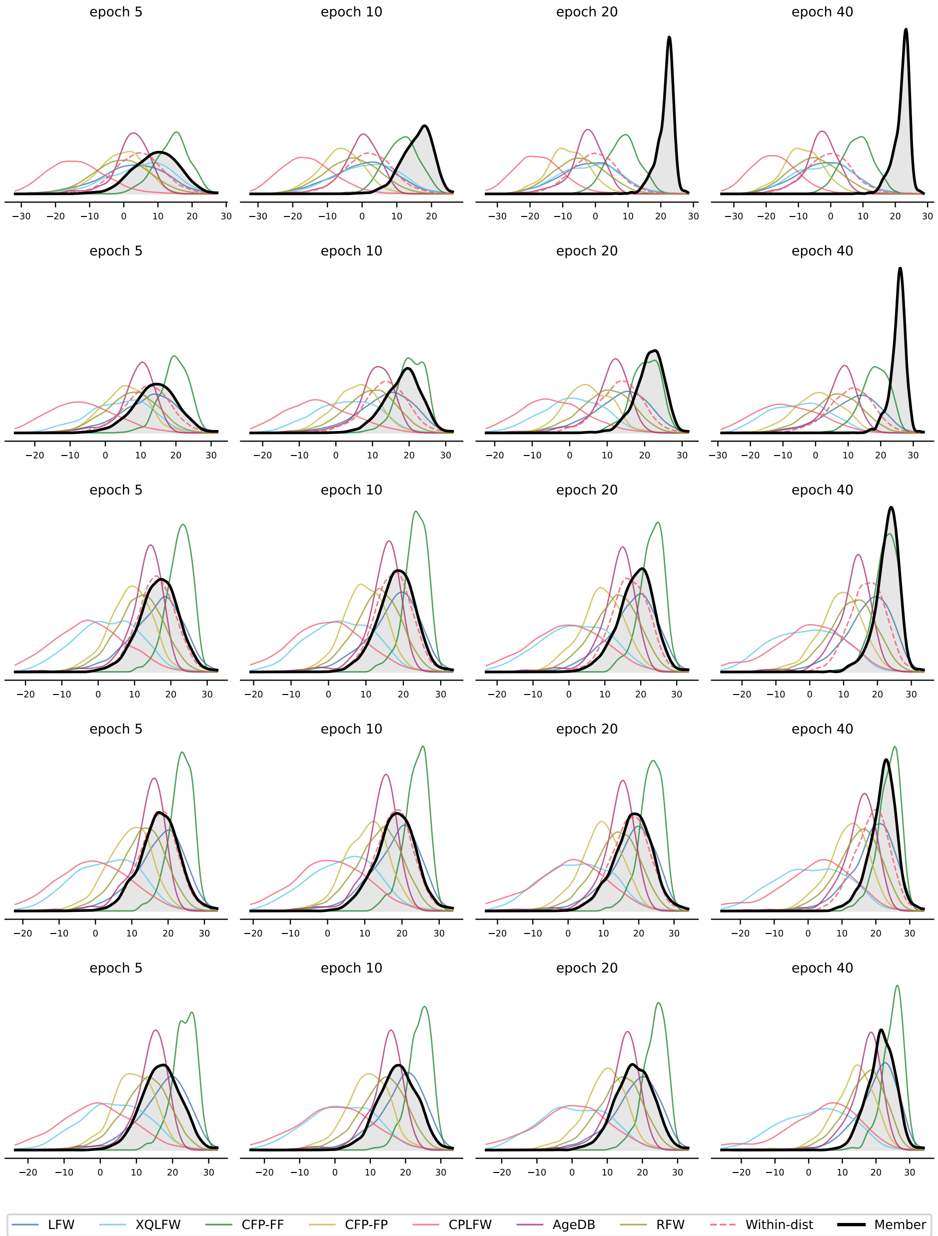


Figure 223. **Vary training duration — IResNet-100 / CosFace.** Rows: $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

6.2. Vary training-set size

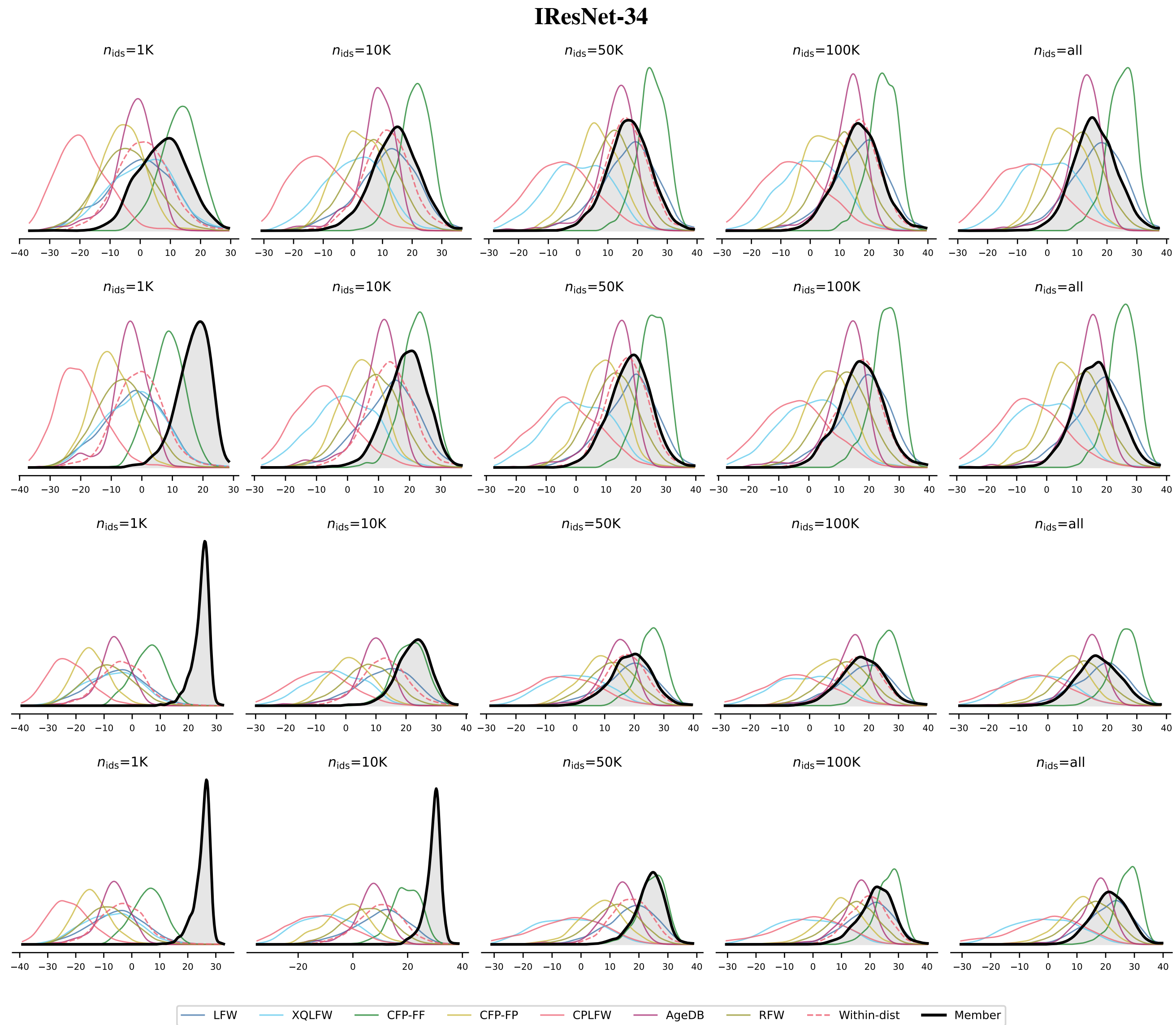


Figure 224. **Vary training-set size — IResNet-34 / ArcFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

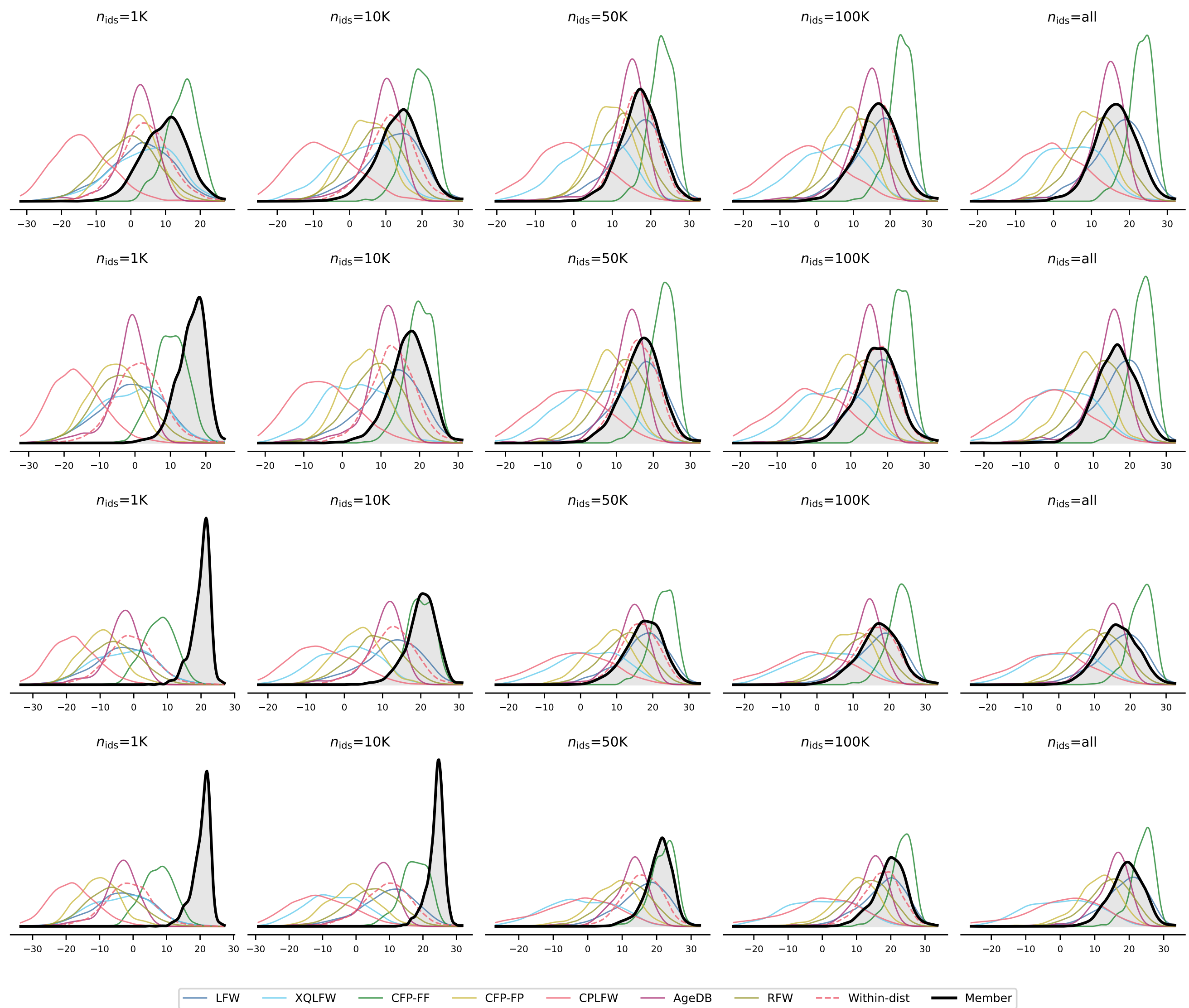


Figure 225. **Vary training-set size — IResNet-34 / CosFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

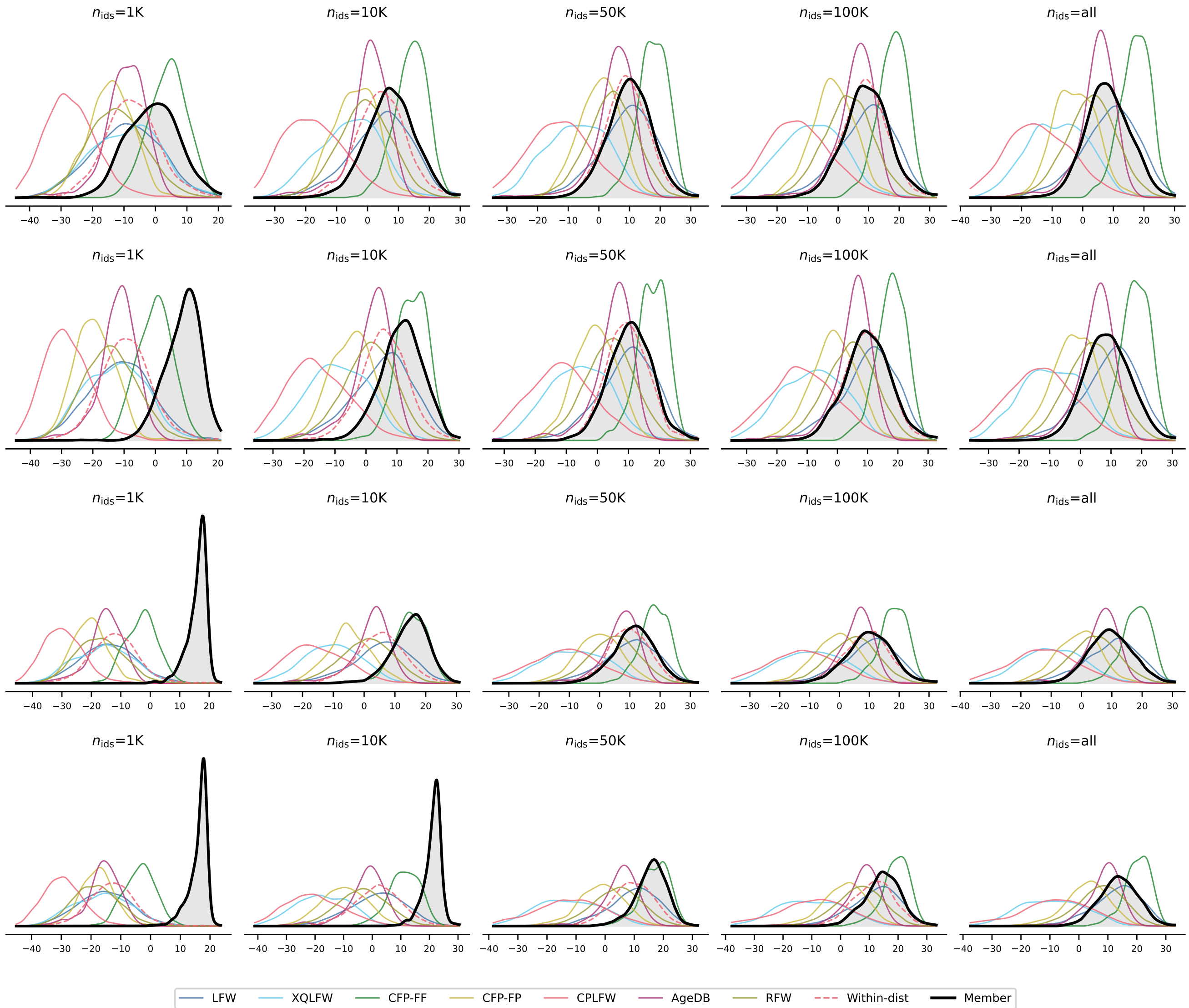


Figure 226. **Vary training-set size — IResNet-34 / MagFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

IResNet-50

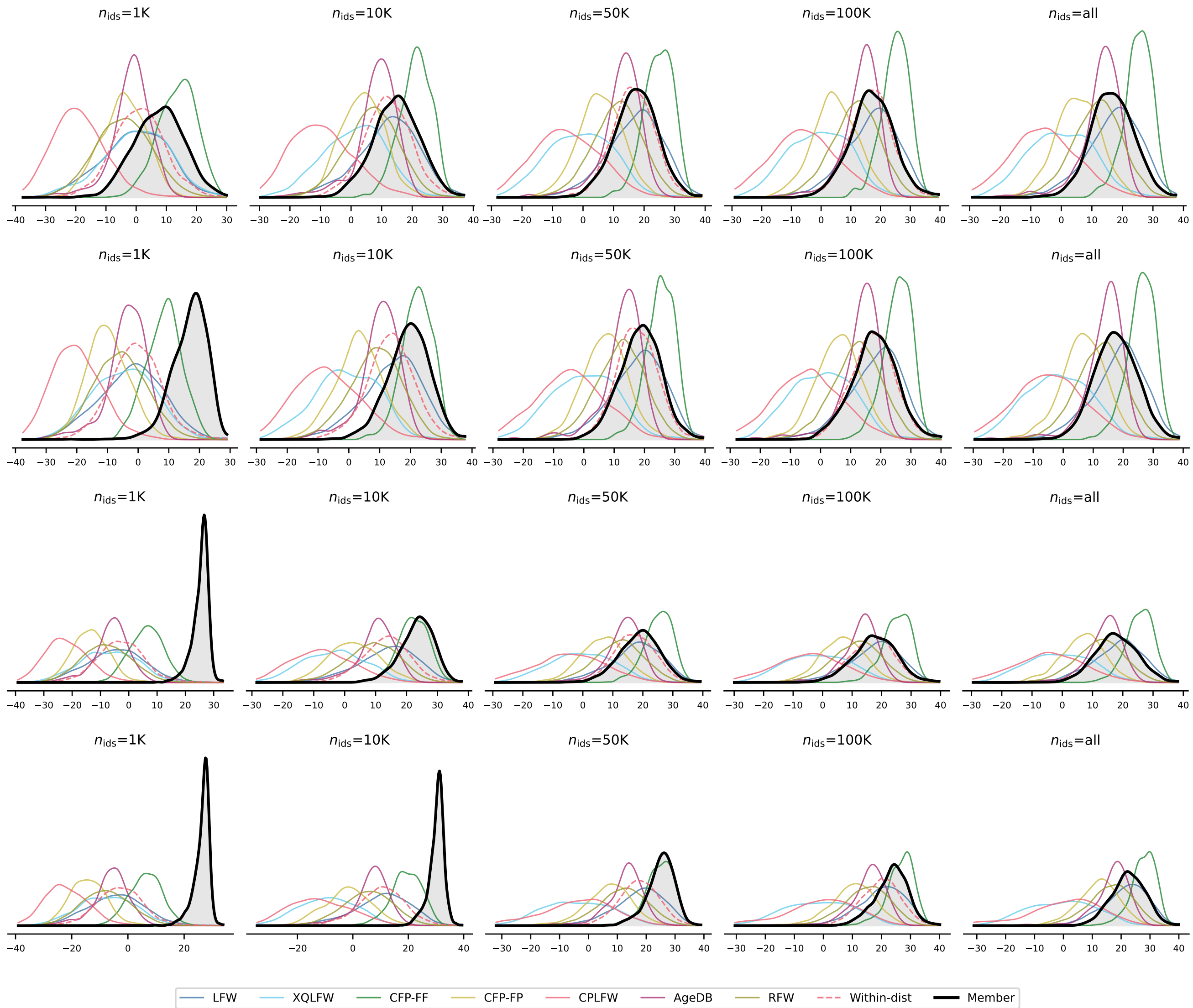


Figure 227. **Vary training-set size — IResNet-50 / ArcFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

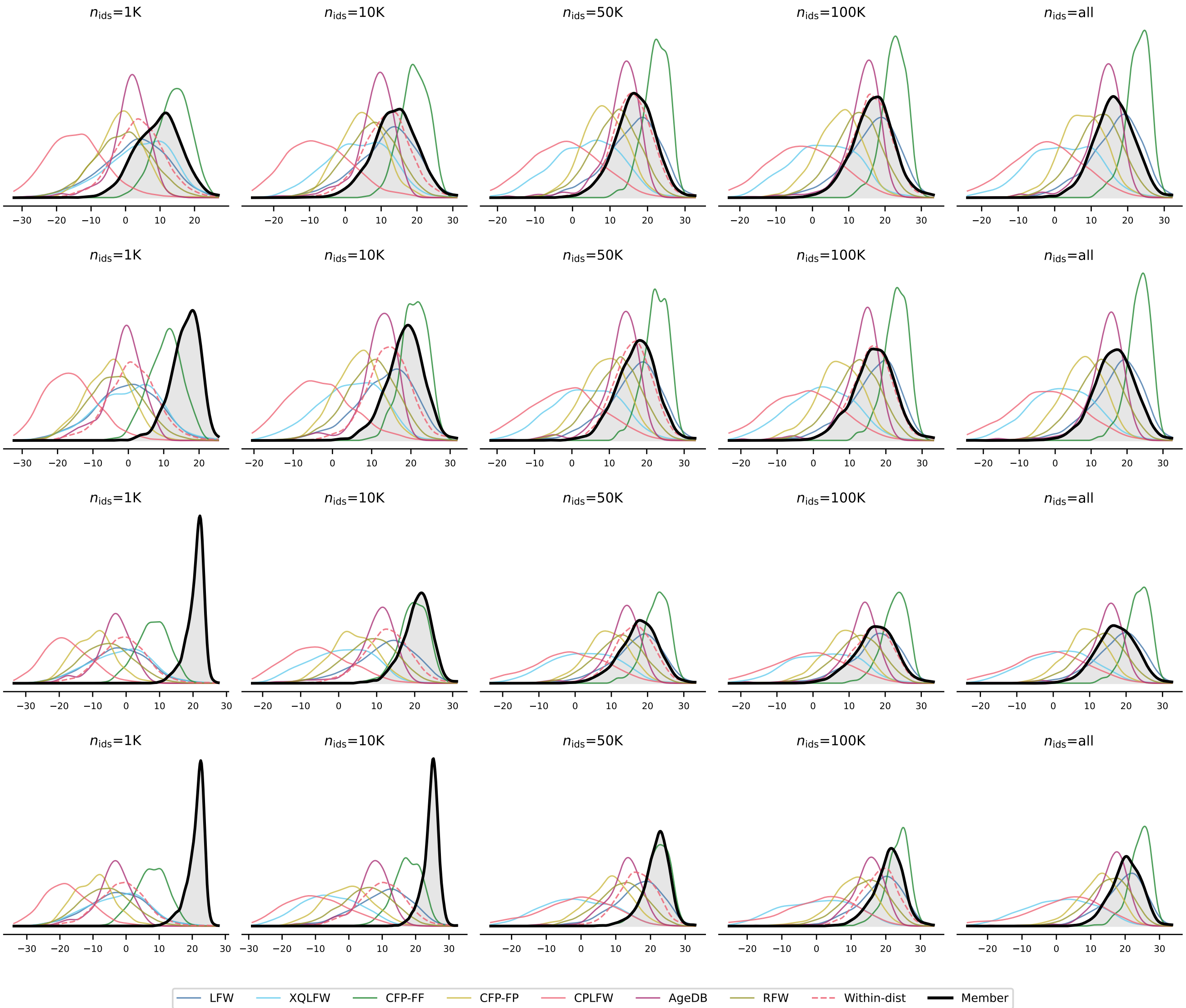


Figure 228. **Vary training-set size — IResNet-50 / CosFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

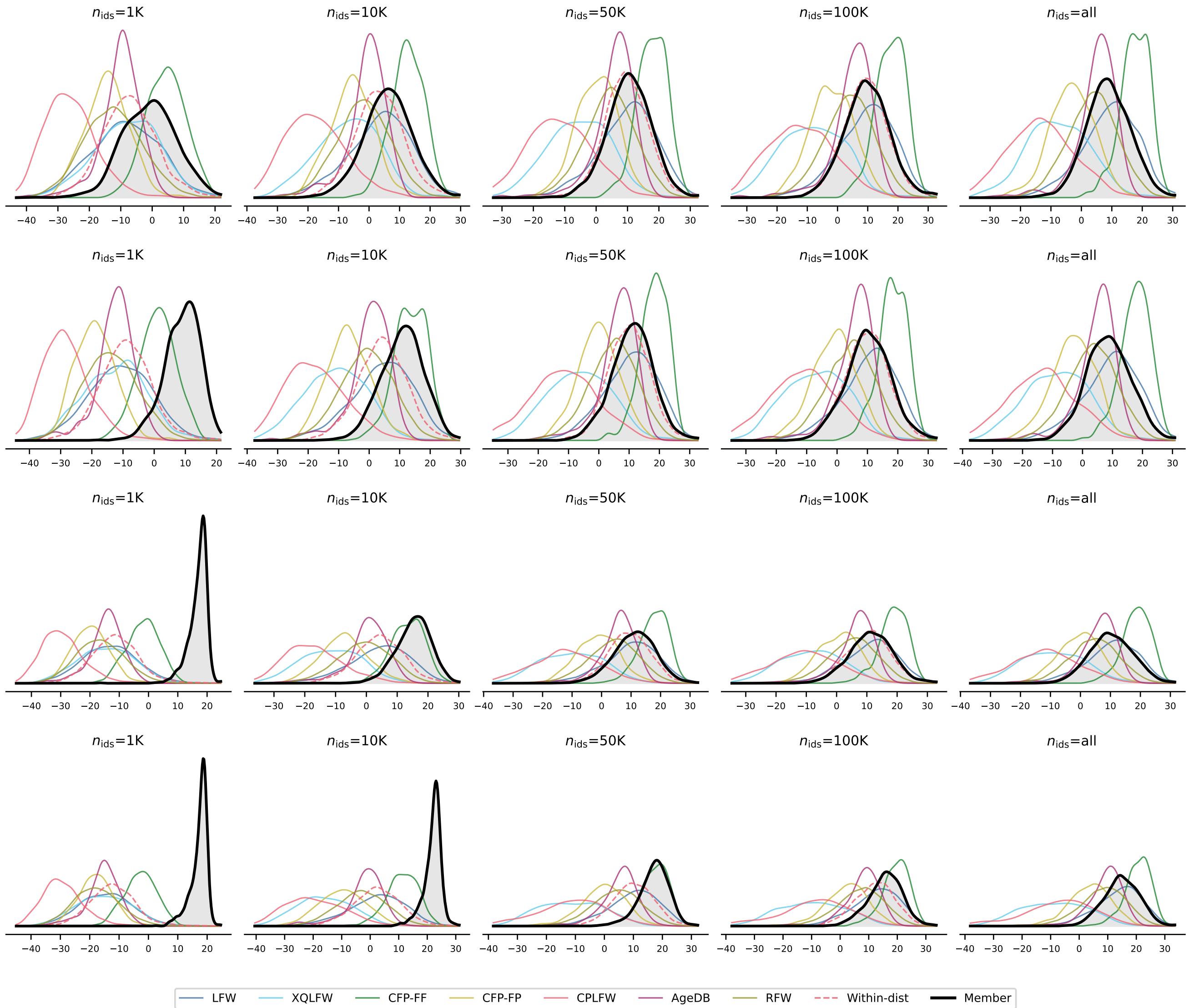


Figure 229. **Vary training-set size — IResNet-50 / MagFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

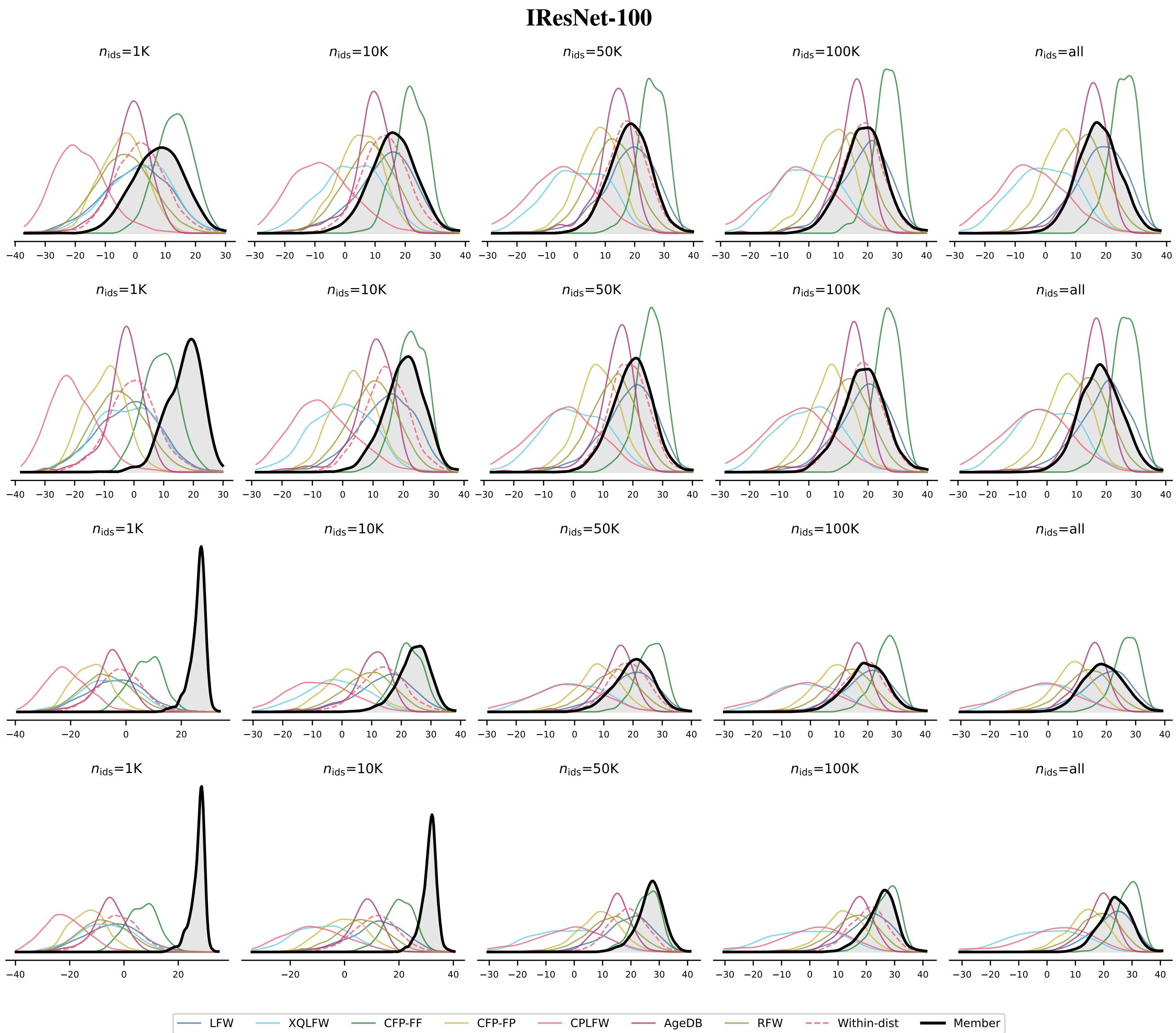


Figure 230. **Vary training-set size — IResNet-100 / ArcFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

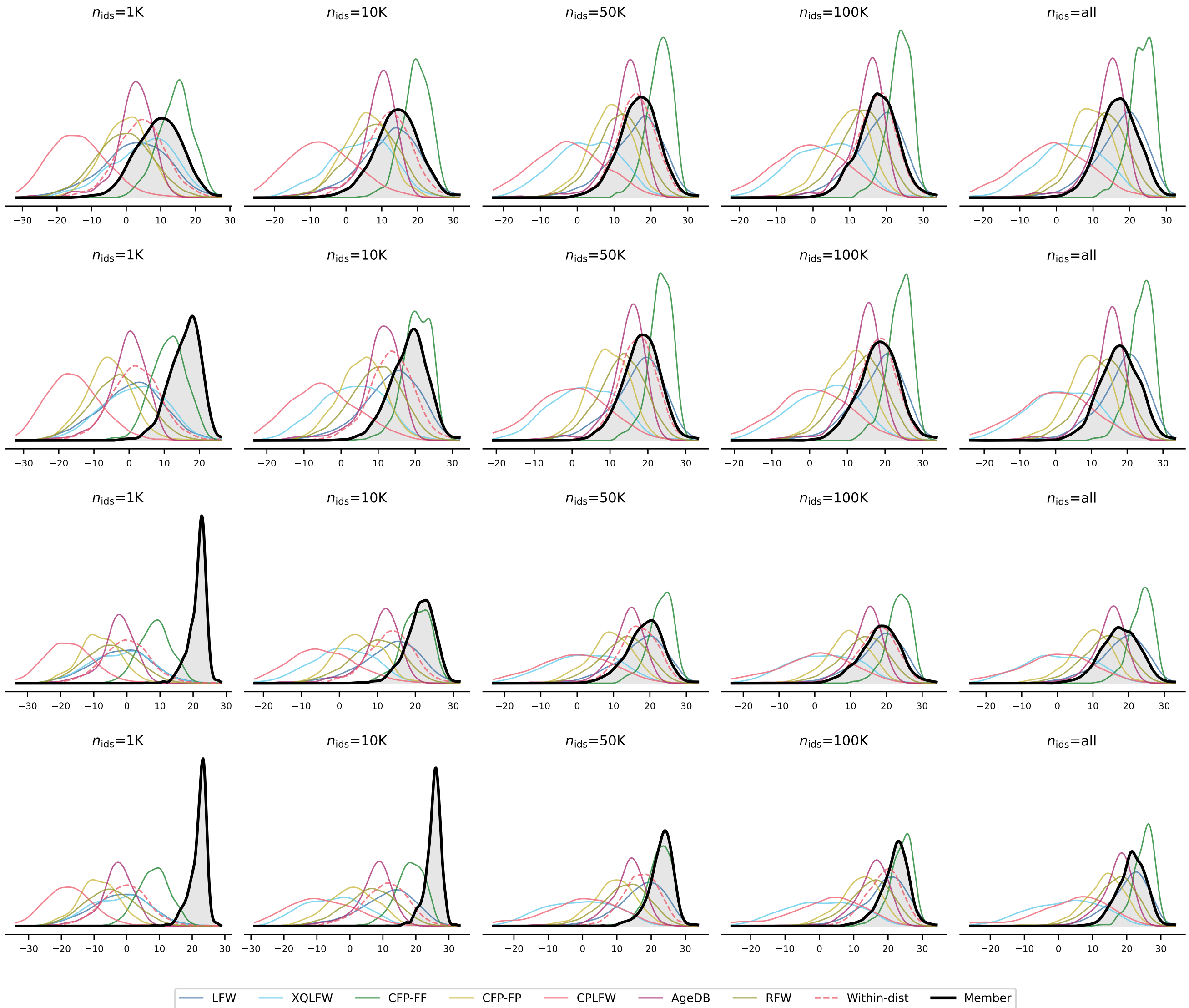


Figure 231. **Vary training-set size — IResNet-100 / CosFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

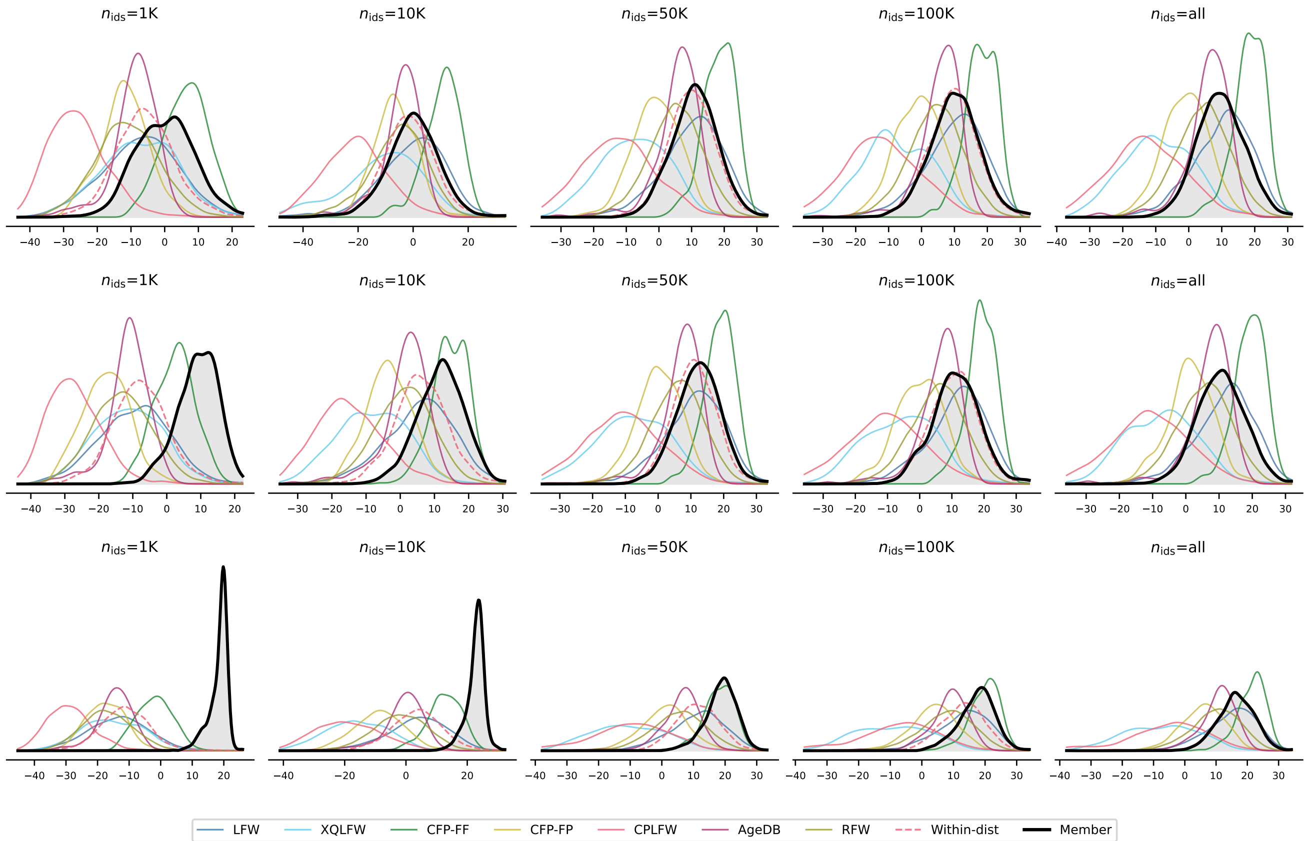


Figure 232. **Vary training-set size — IResNet-100 / MagFace.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

6.3. Vary loss head

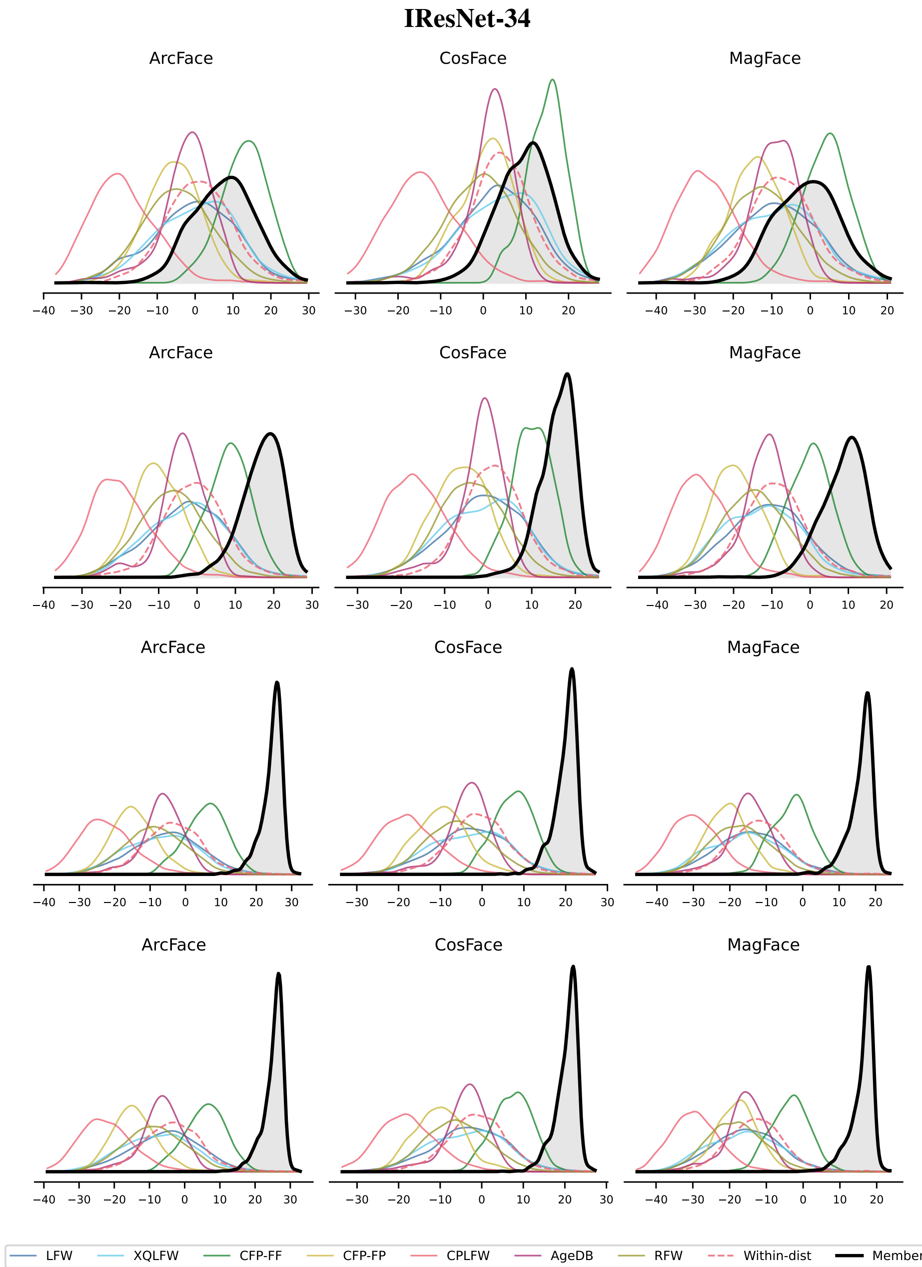


Figure 233. **Vary loss head — IResNet-34 /** $n_{\textbf{ids}} = 1K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

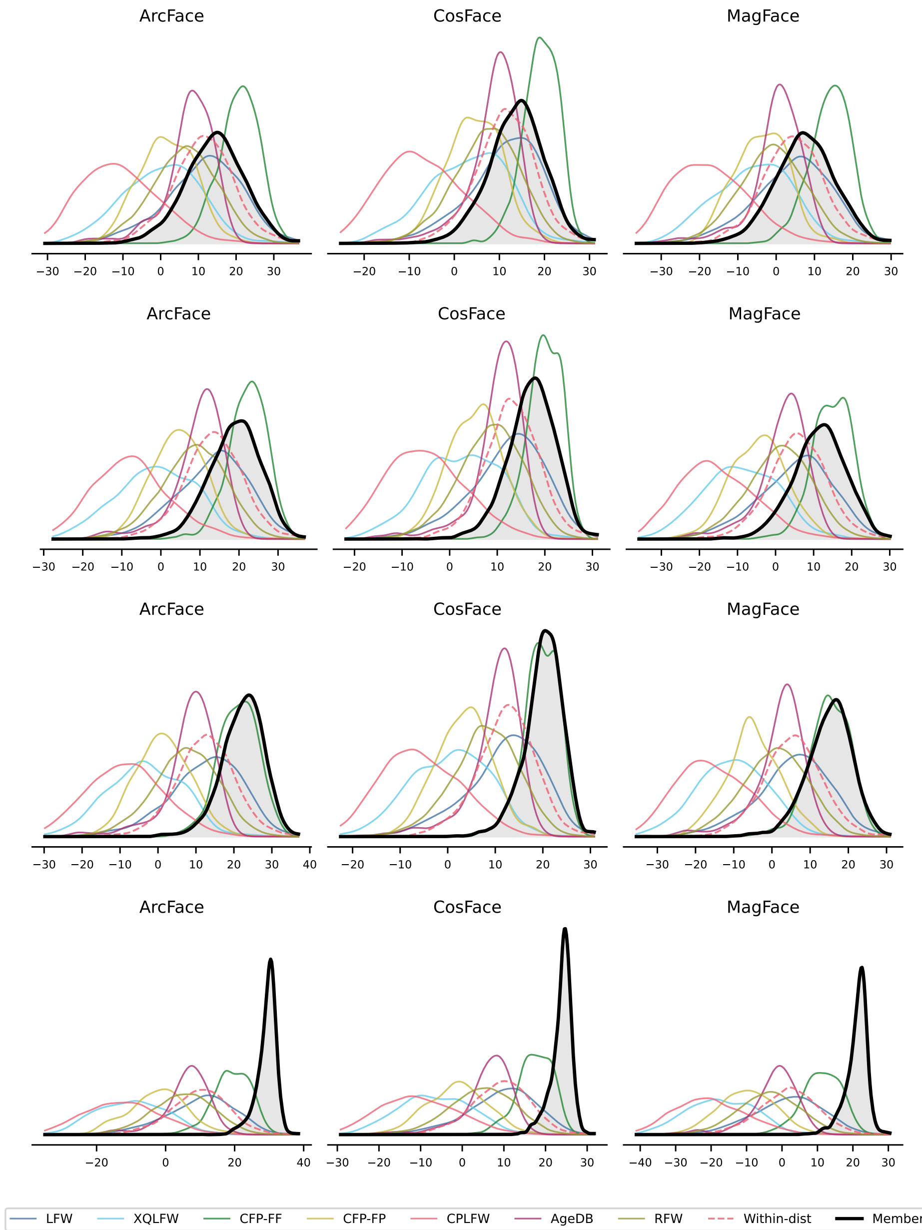


Figure 234. **Vary loss head — IResNet-34 / $n_{\text{ids}} = 10K$.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

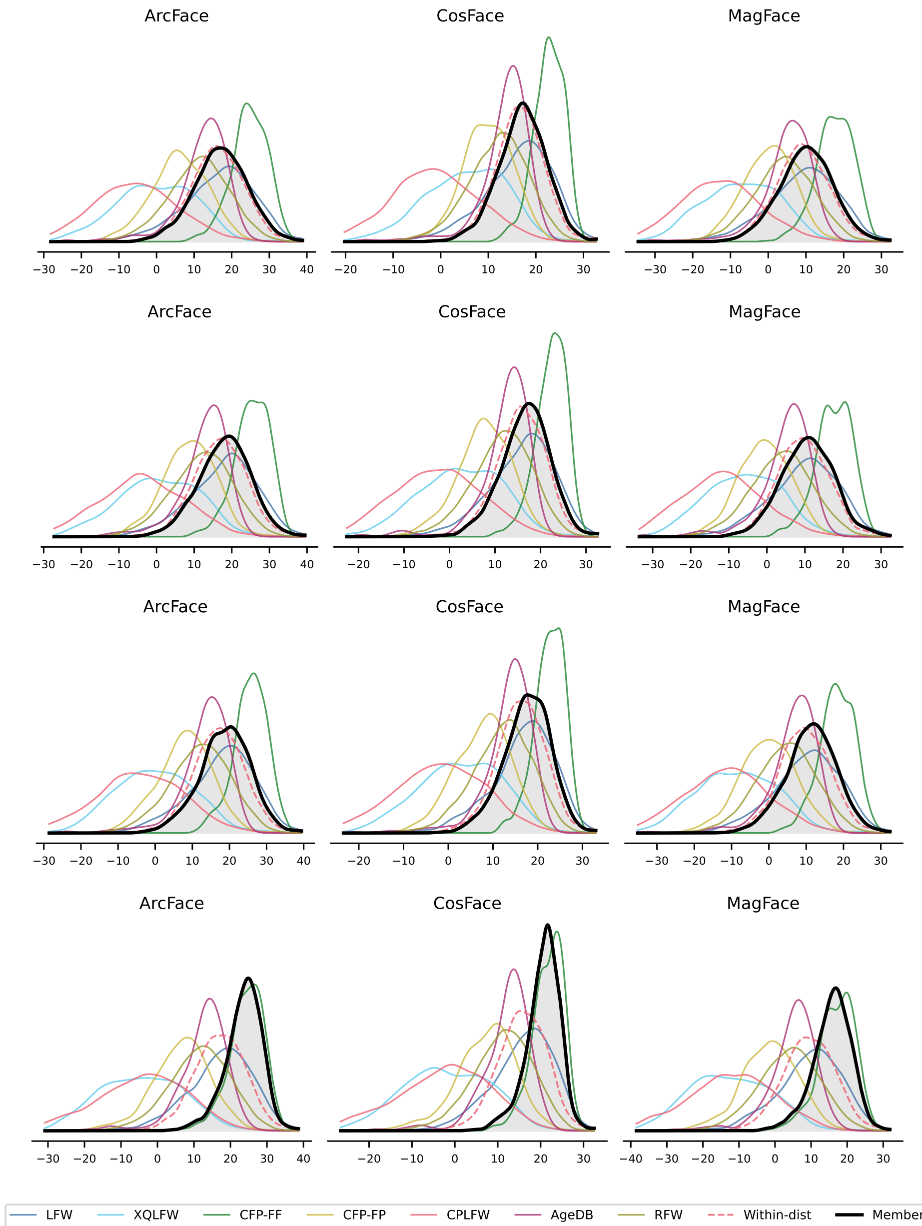


Figure 235. **Vary loss head — IResNet-34 /** $n_{\mathbf{ids}} = 50K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

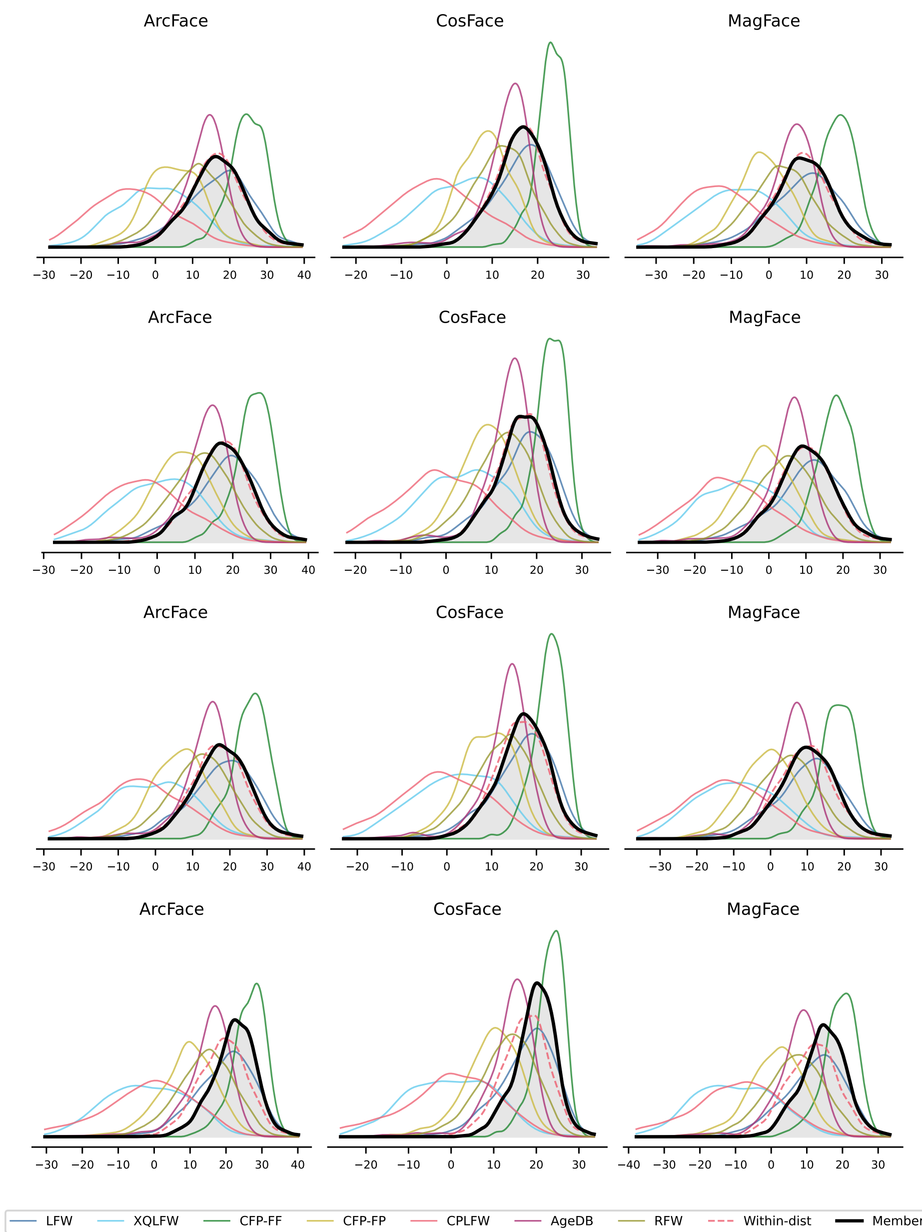


Figure 236. **Vary loss head — IResNet-34 /** $n_{\text{ids}} = 100K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

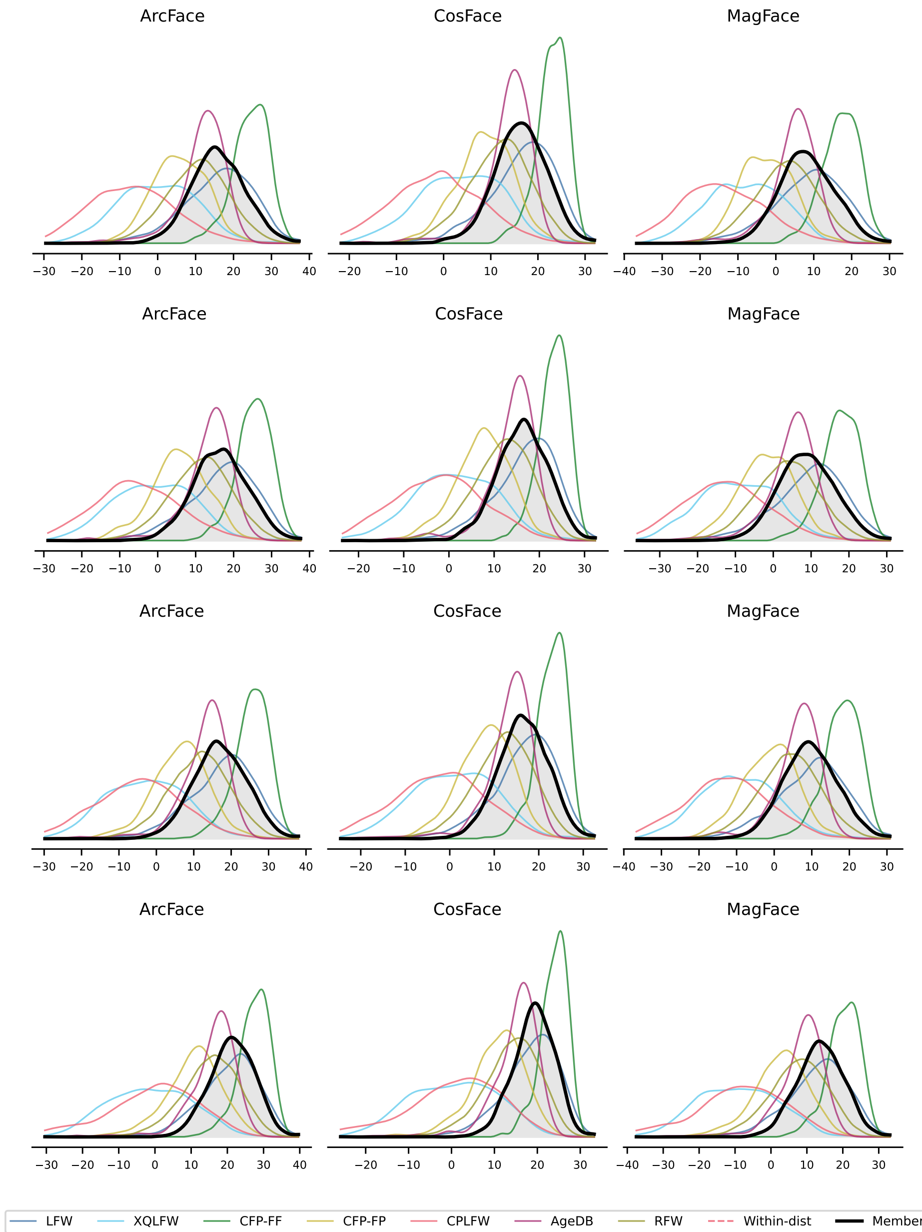


Figure 237. **Vary loss head — IResNet-34 /** $n_{\text{ids}} = All$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

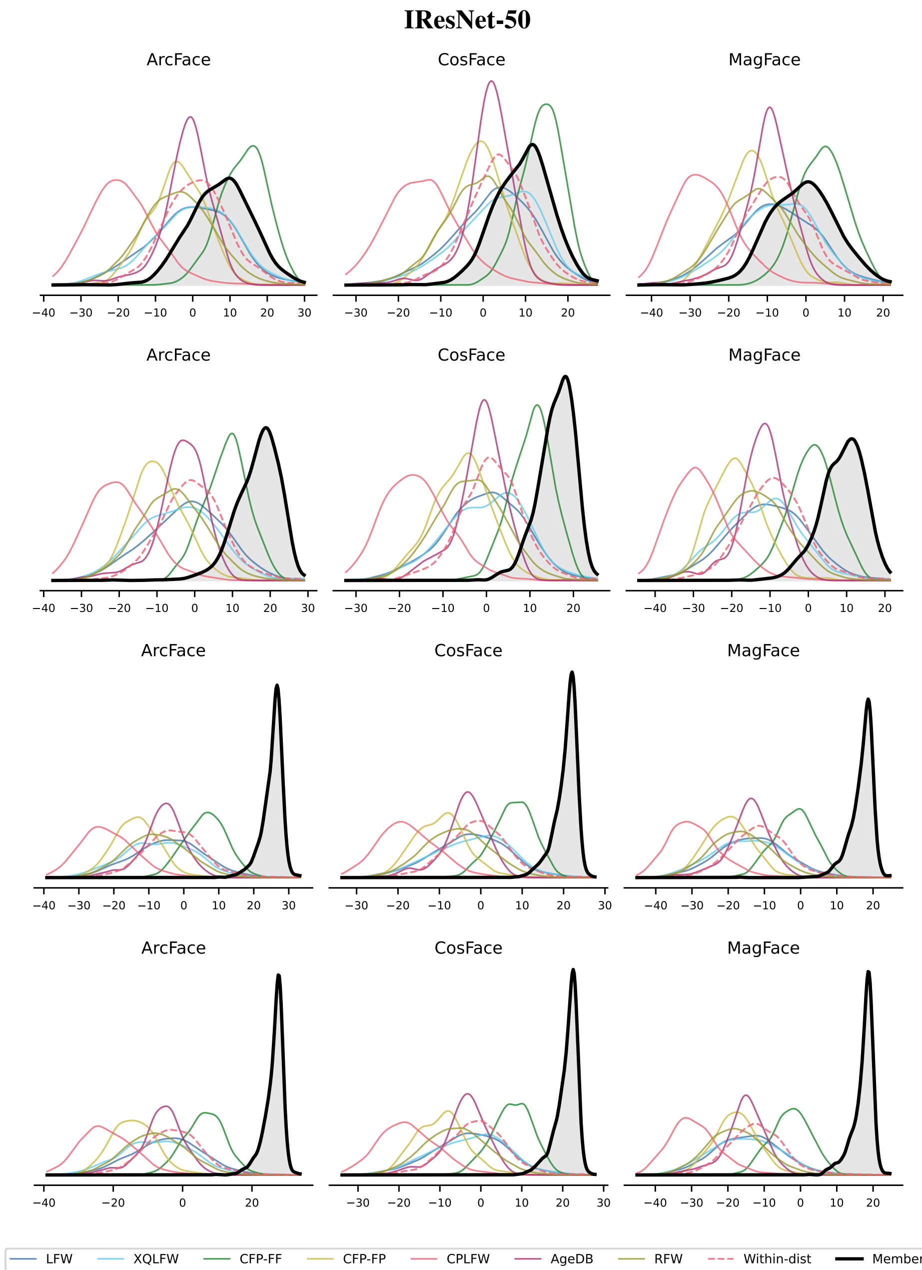


Figure 238. **Vary loss head — IResNet-50 / $n_{\text{ids}} = 1K$.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

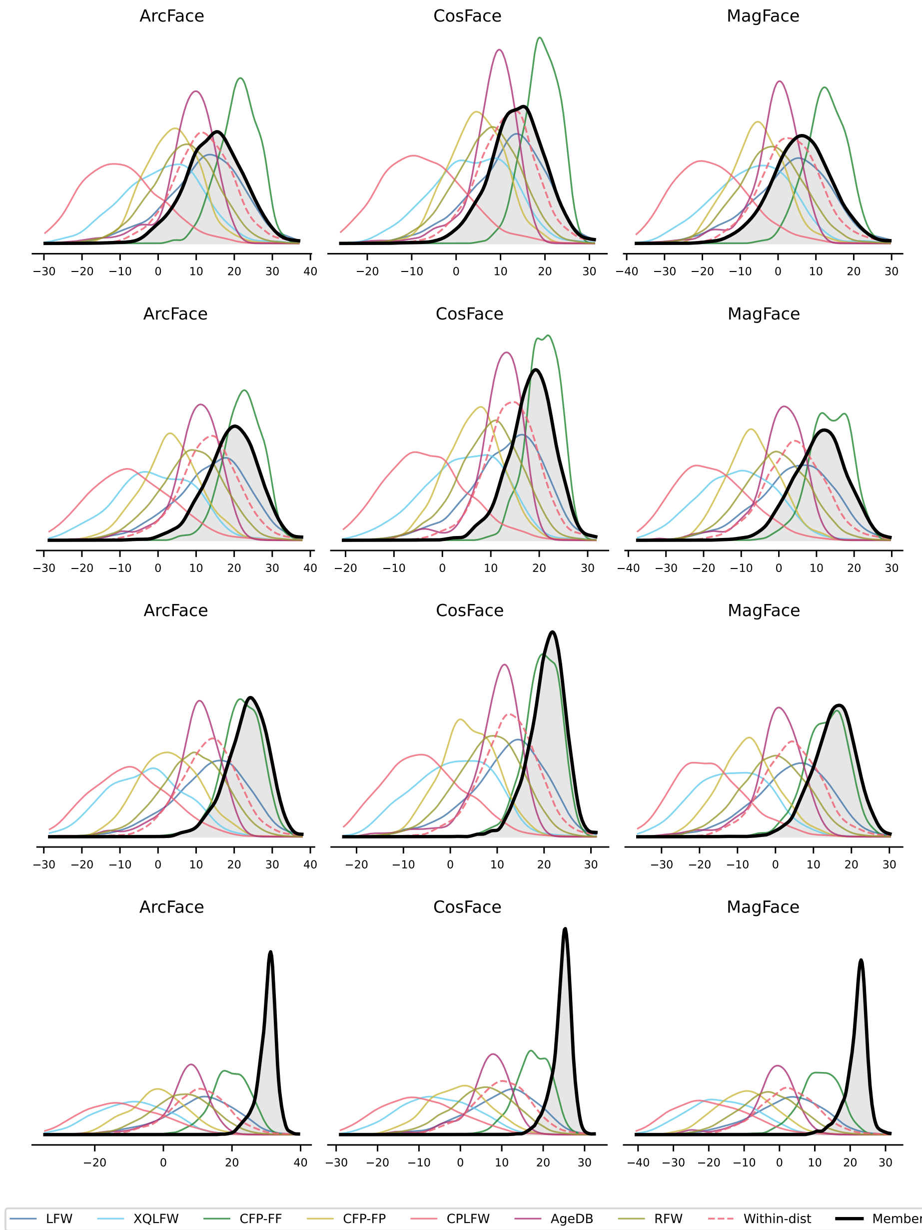


Figure 239. **Vary loss head — IResNet-50 /** $n_{\text{ids}} = 10K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

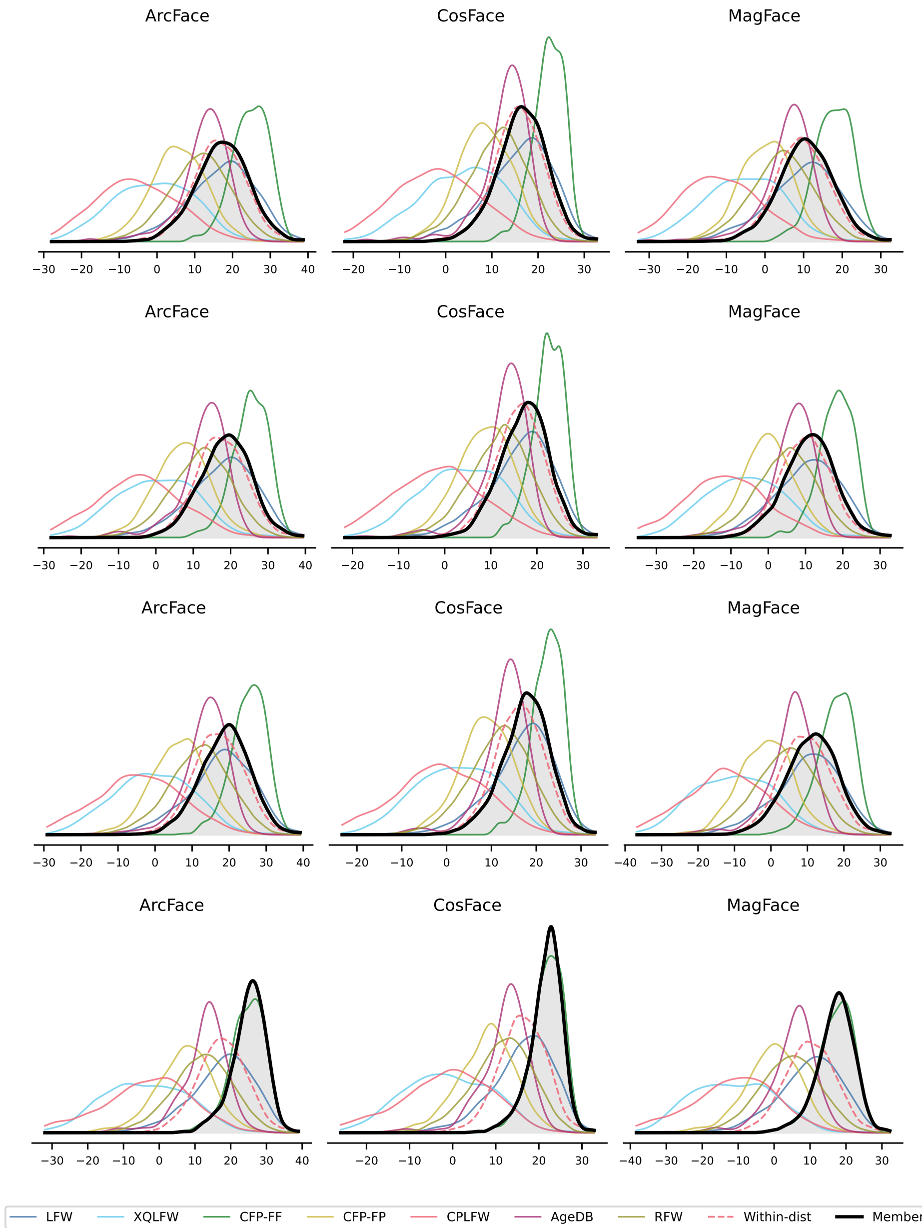


Figure 240. **Vary loss head — IResNet-50 /** $n_{\text{ids}} = 50K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

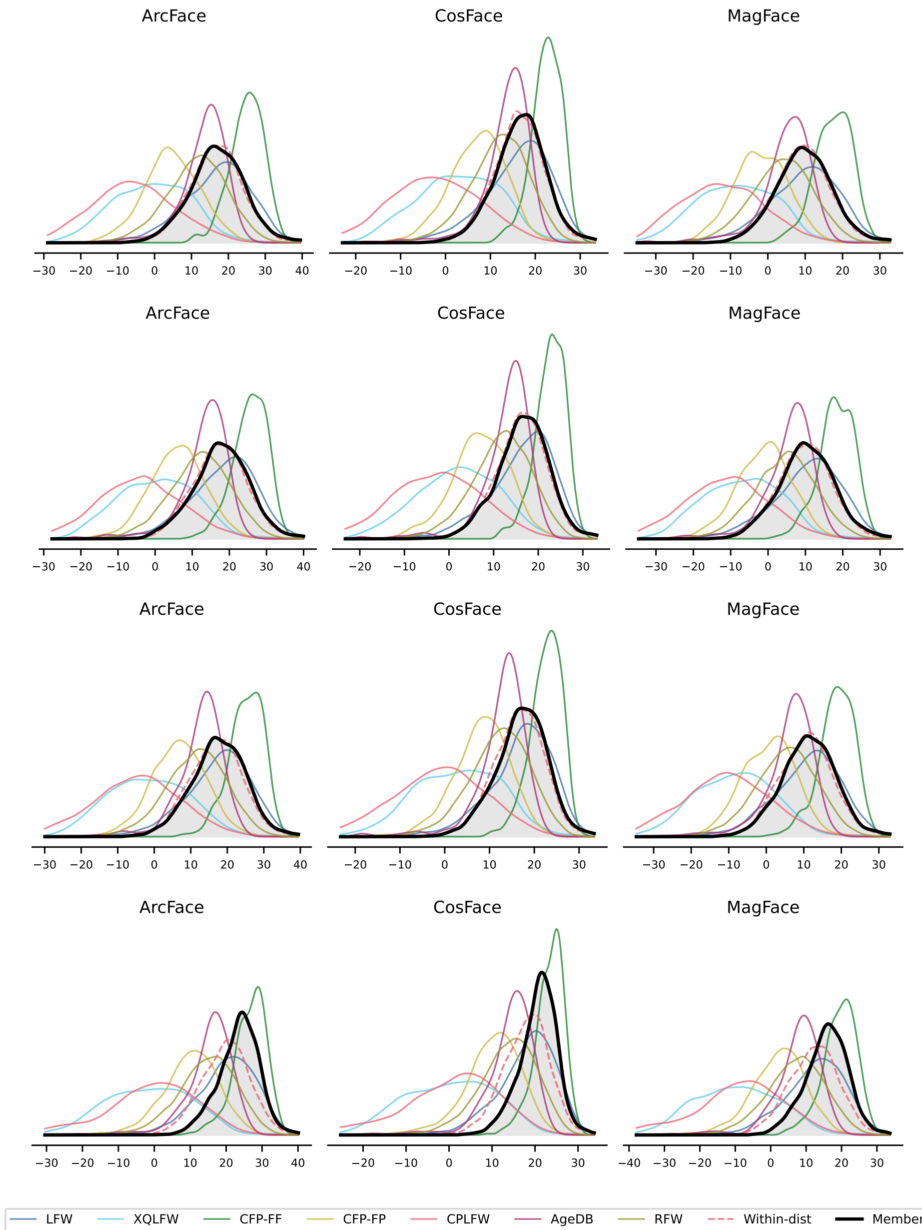


Figure 241. **Vary loss head — IResNet-50 / $n_{\text{ids}} = 100K$.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

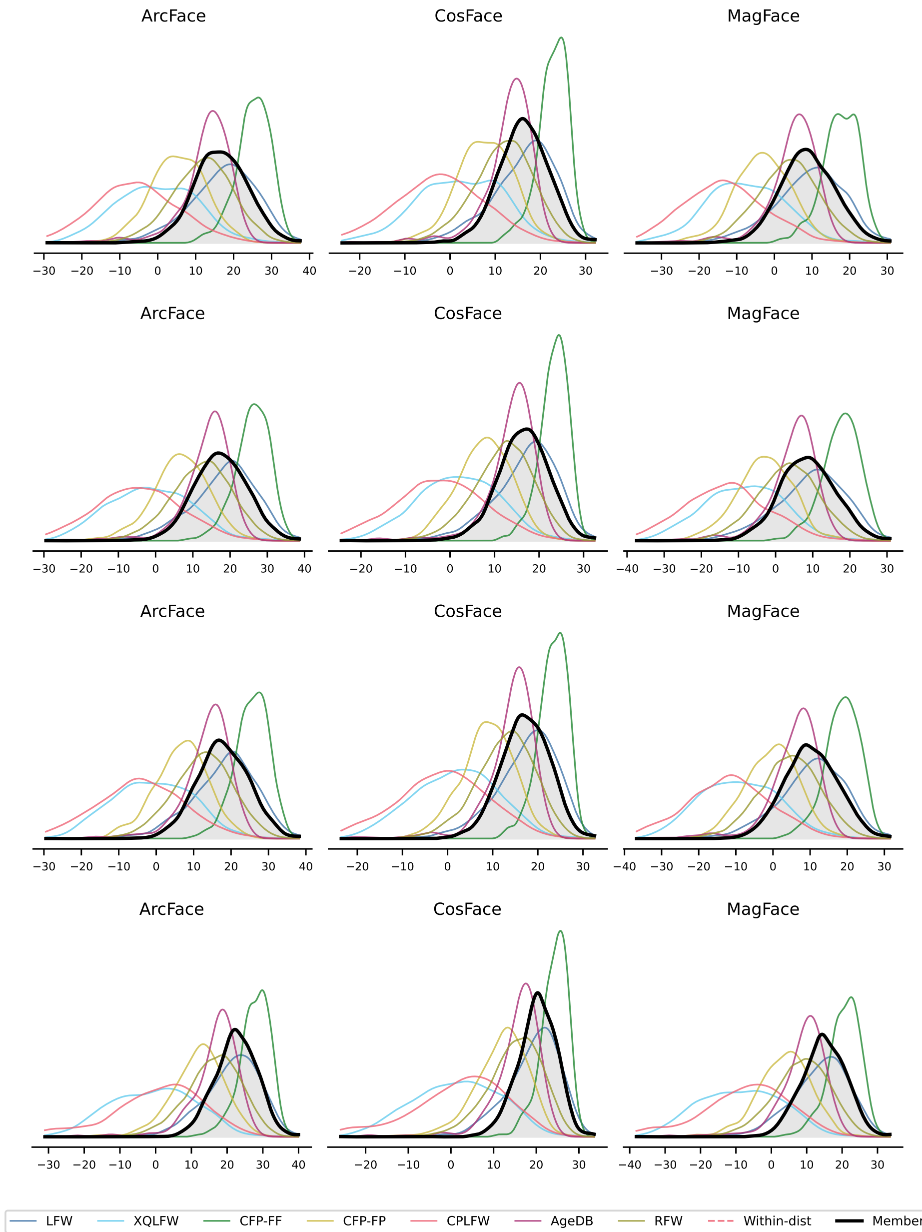


Figure 242. **Vary loss head — IResNet-50 / $n_{\text{ids}} = All$.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

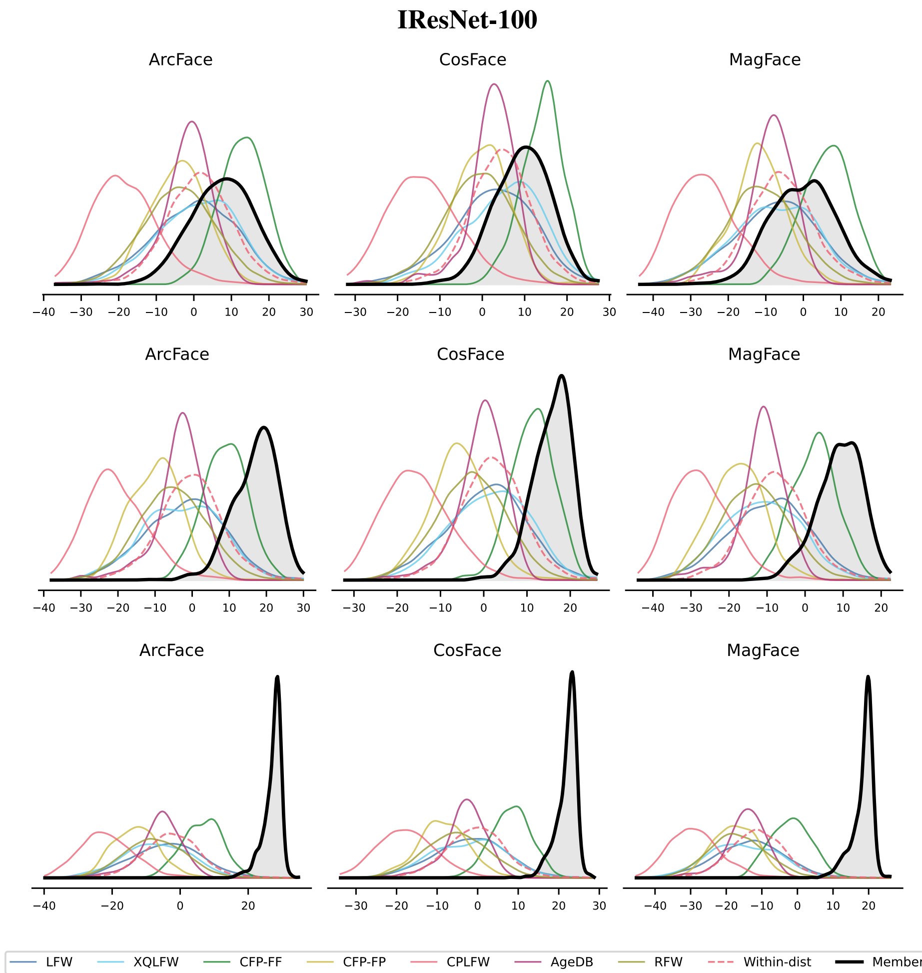


Figure 243. **Vary loss head — IResNet-100 /** $n_{\text{ids}} = 1K$**.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

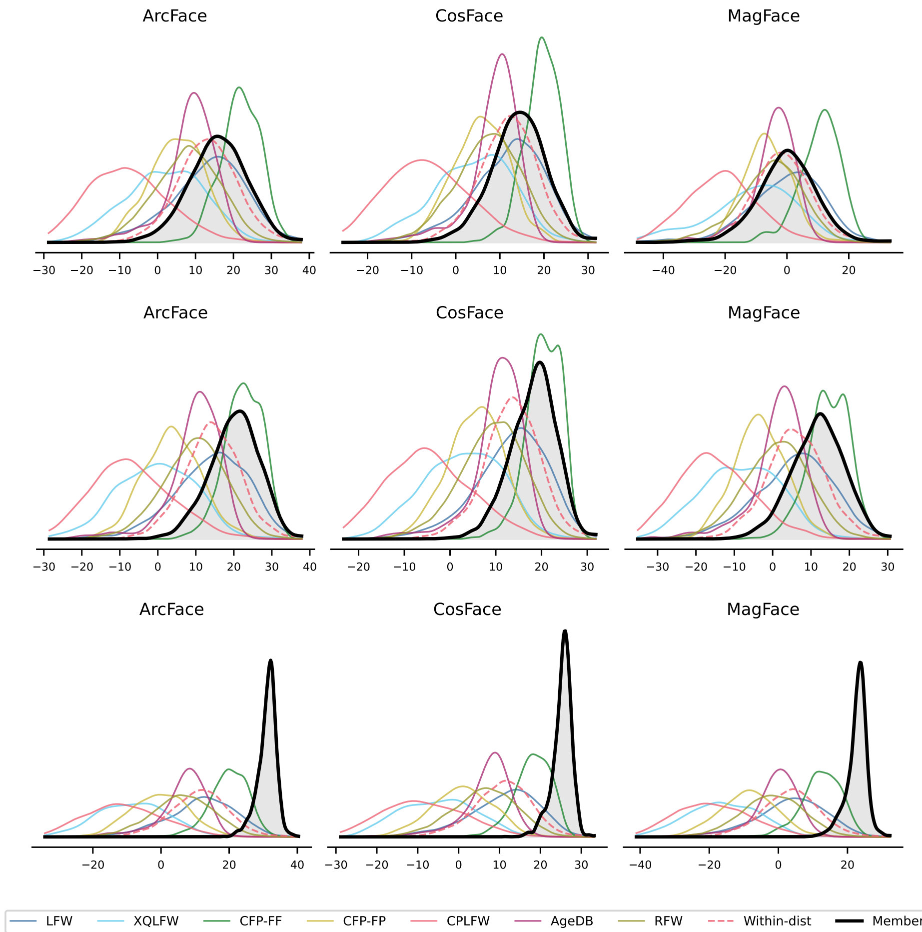


Figure 244. **Vary loss head — IResNet-100 /** $n_{\mathbf{ids}} = 10K$**.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

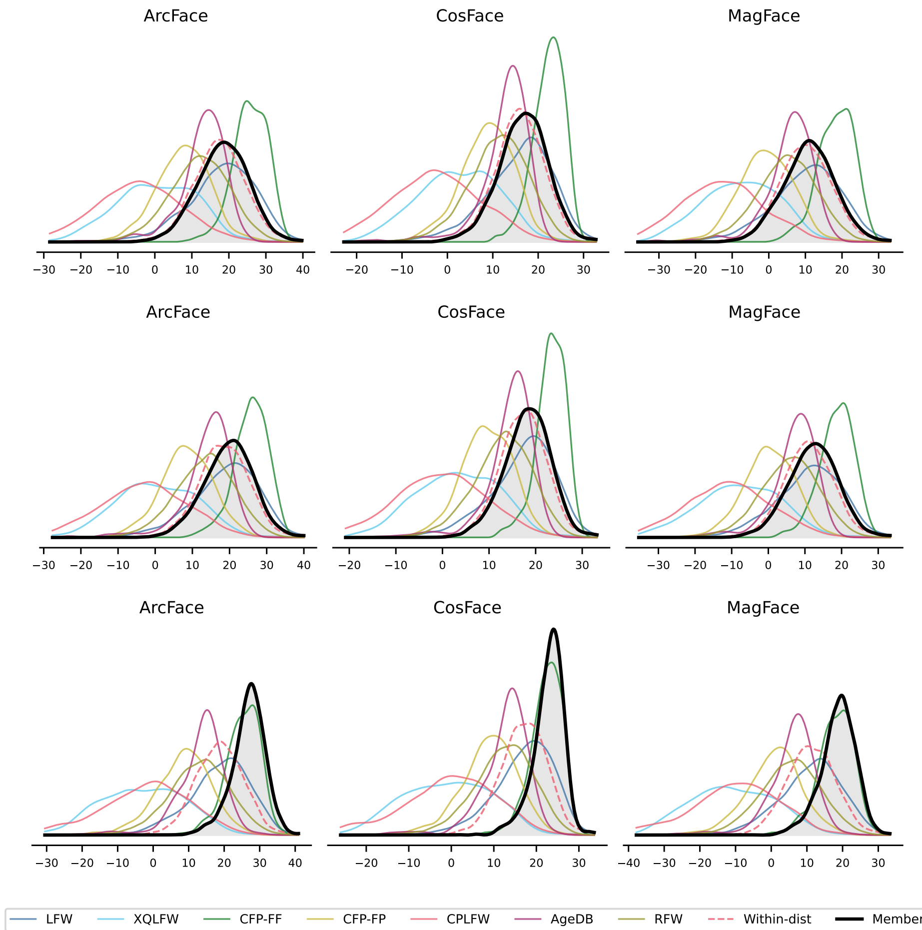


Figure 245. **Vary loss head — IResNet-100 /** $n_{\text{ids}} = 50K$**.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

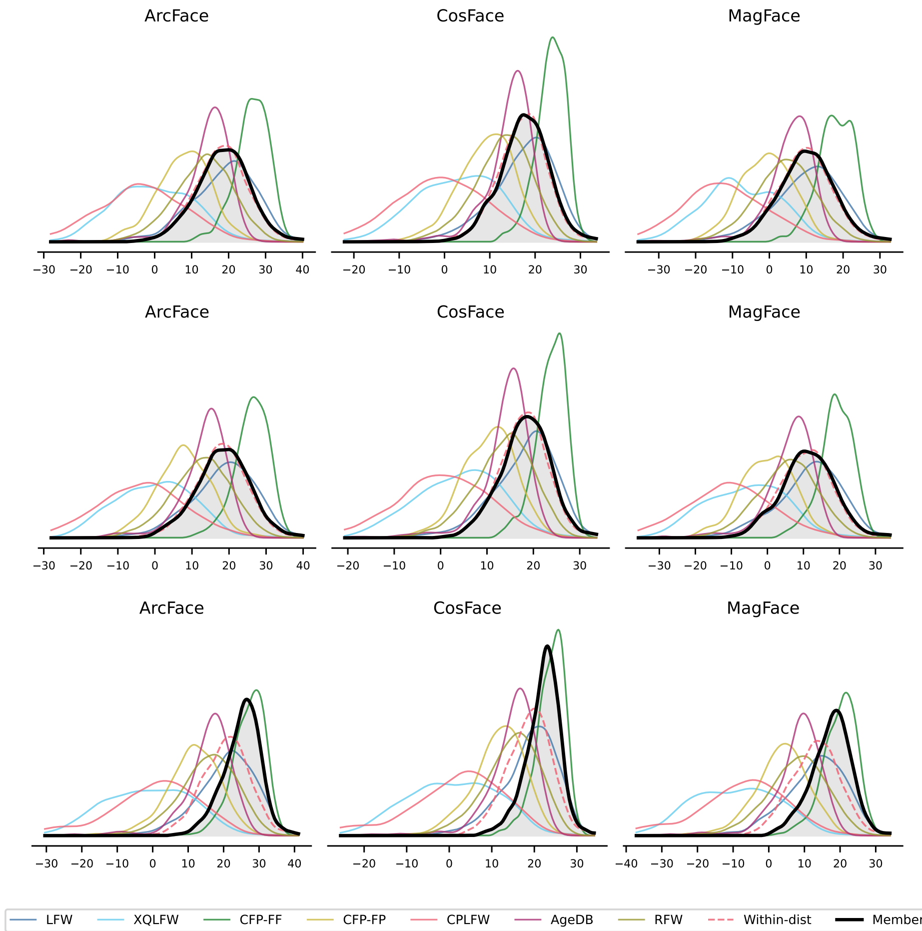


Figure 246. **Vary loss head — IResNet-100 /** $n_{\text{ids}} = 100K$**.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

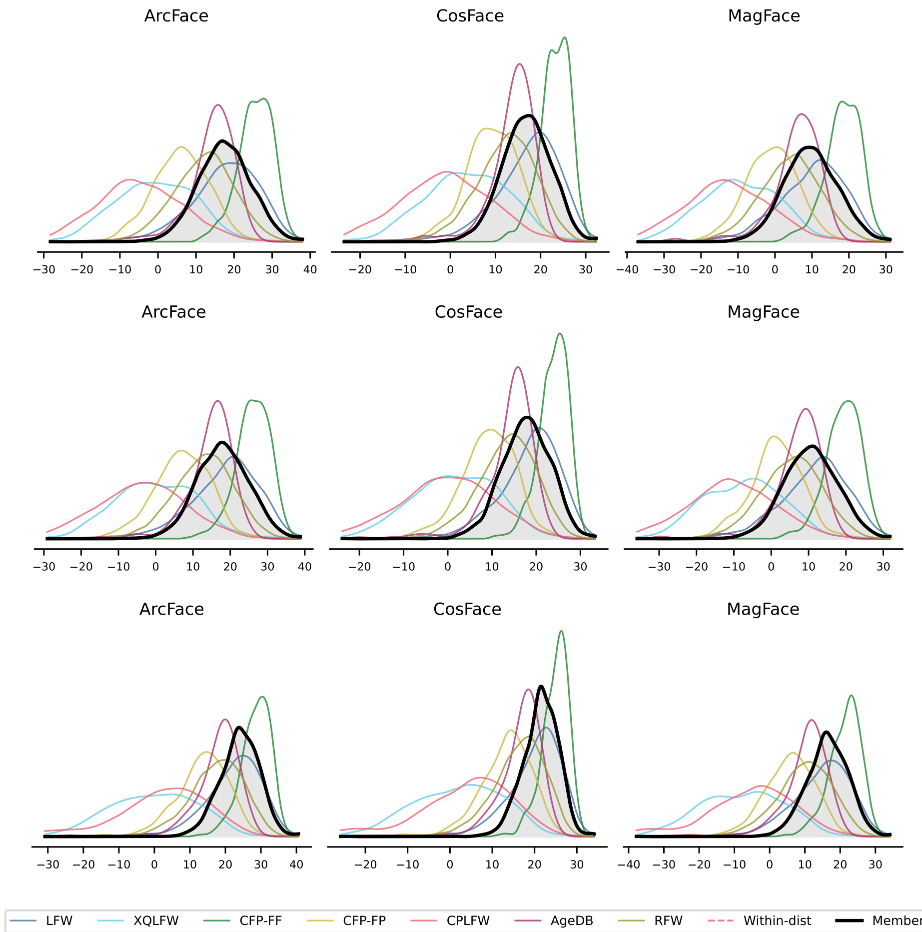


Figure 247. **Vary loss head — IResNet-100 /** $n_{\text{ids}} = All$**.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

6.4. Vary backbone

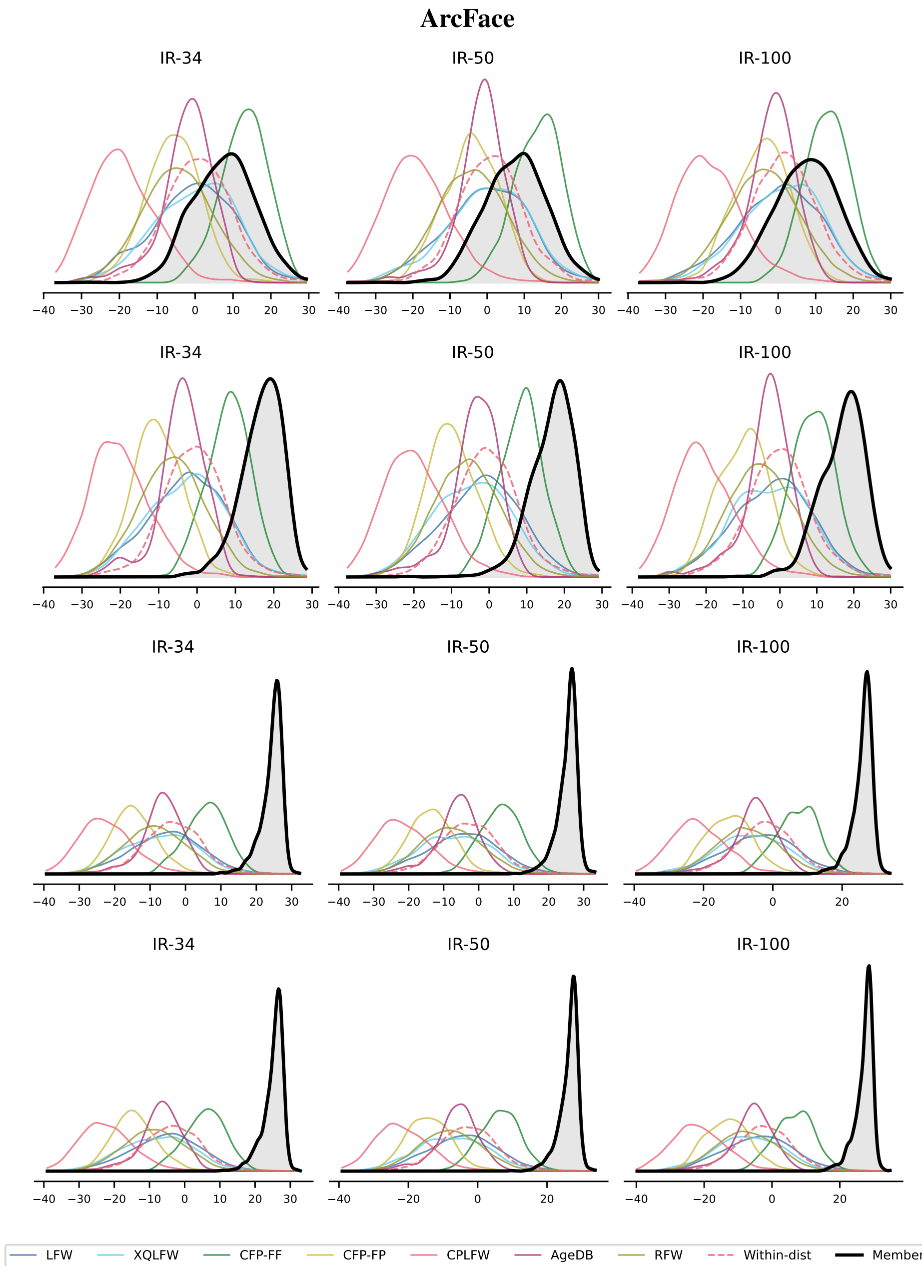


Figure 248. **Vary backbone — ArcFace /** $n_{\mathbf{ids}} = 1K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

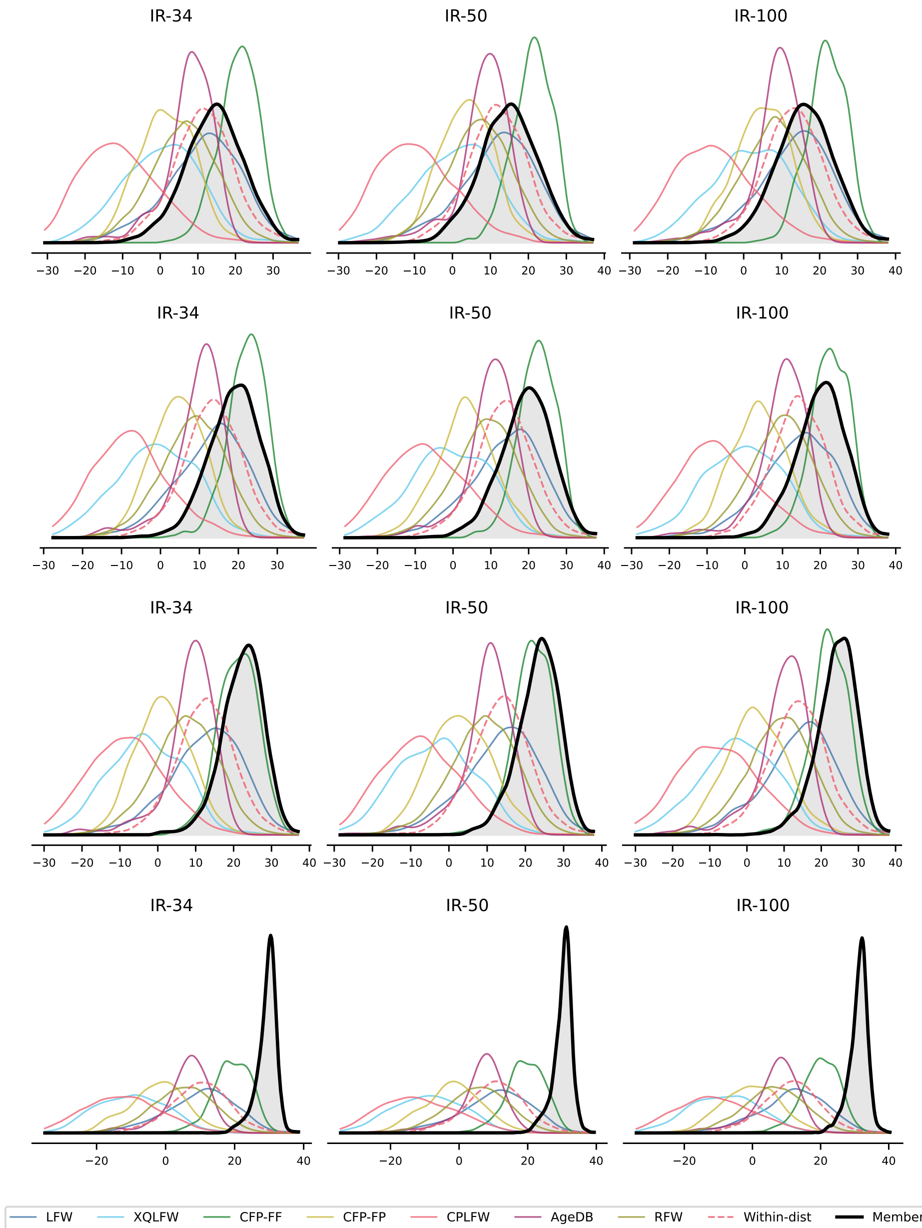


Figure 249. **Vary backbone — ArcFace /** $n_{\text{ids}} = 10K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

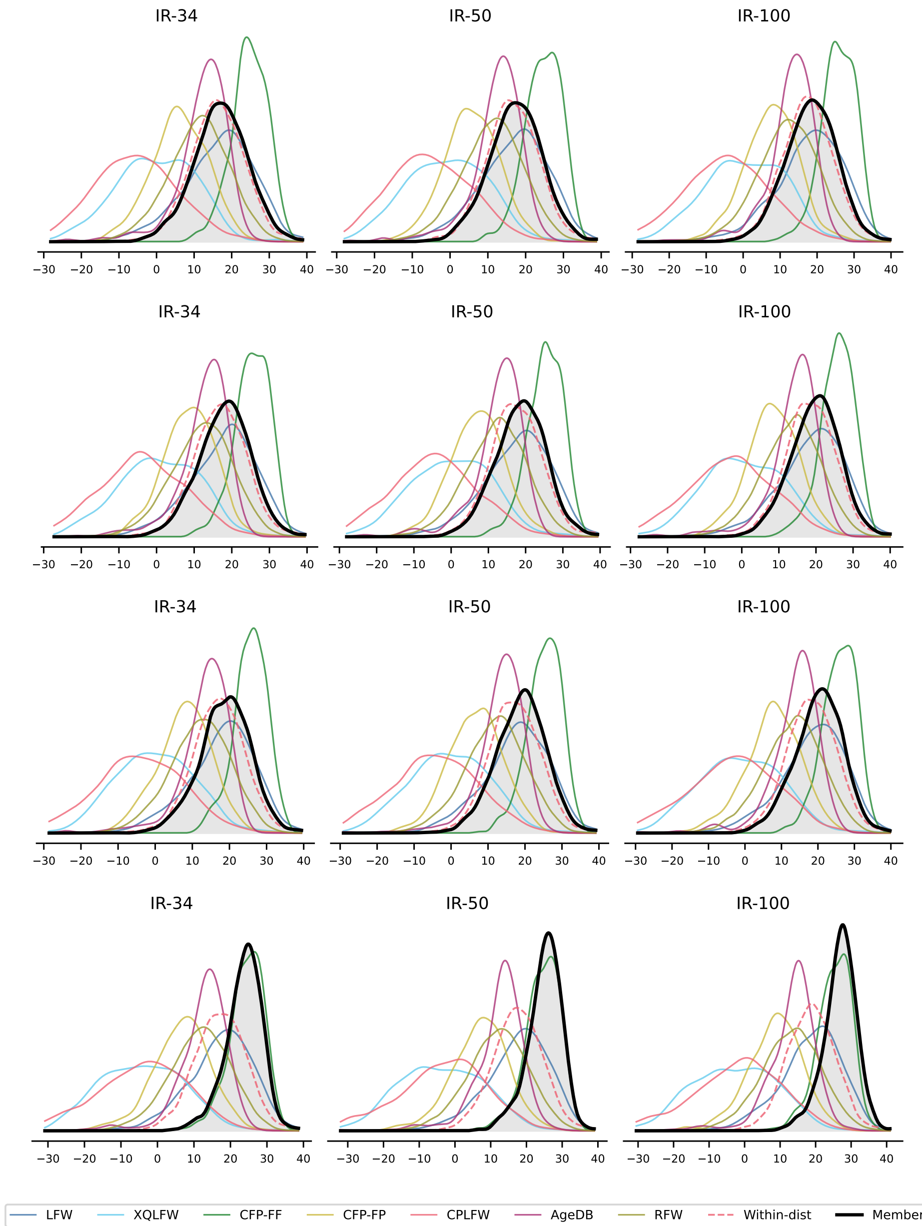


Figure 250. **Vary backbone — ArcFace /** $n_{\text{ids}} = 50K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

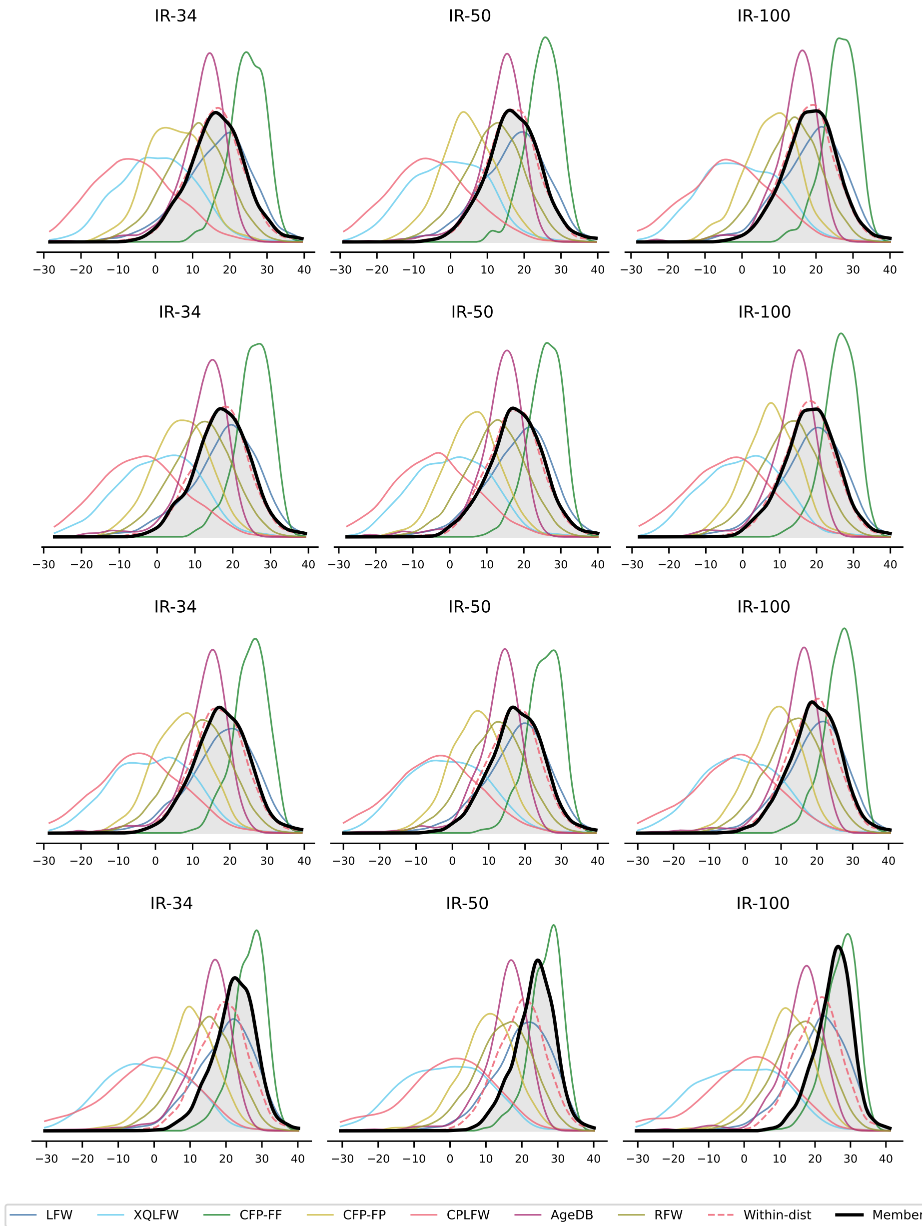


Figure 251. **Vary backbone — ArcFace /** $n_{\text{ids}} = 100K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

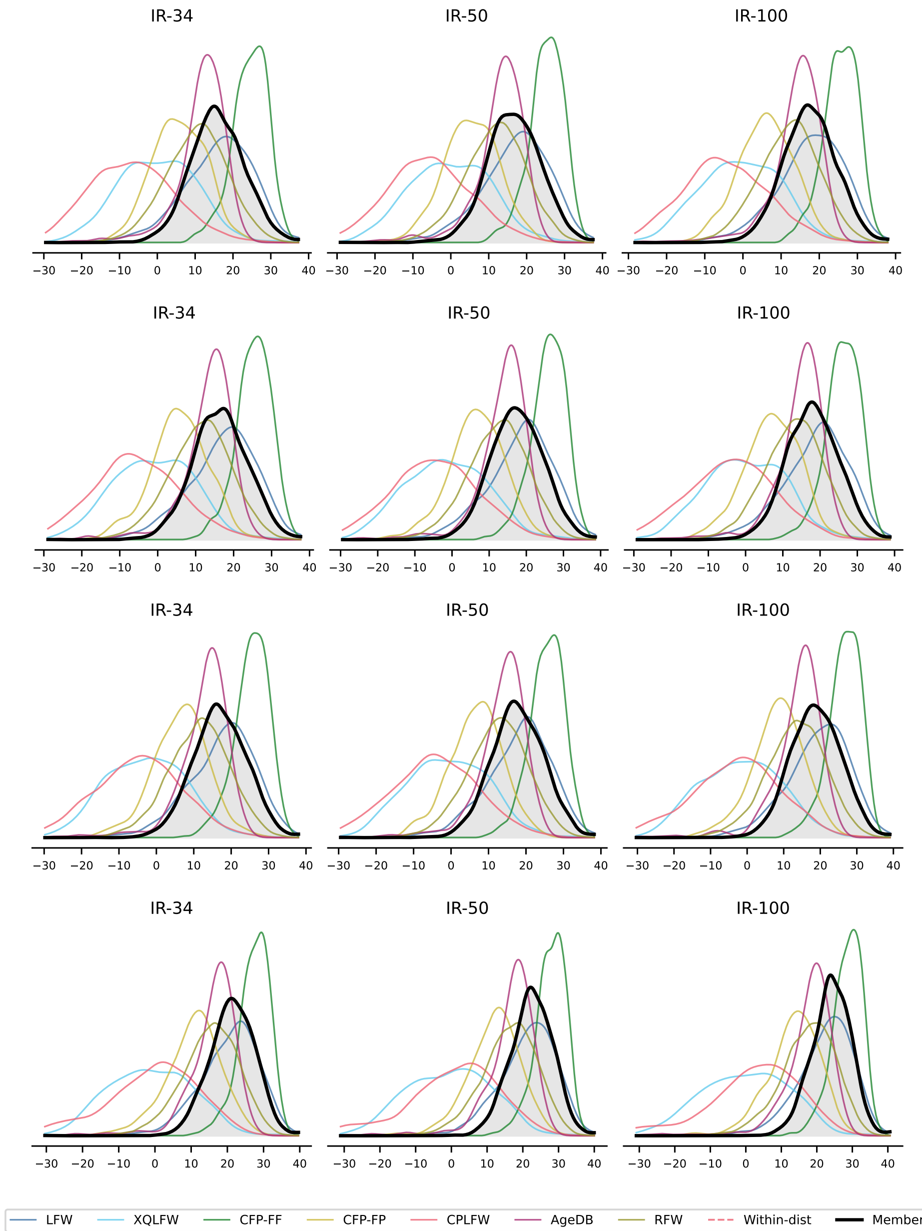


Figure 252. **Vary backbone — ArcFace /** $n_{\mathbf{ids}} = All$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

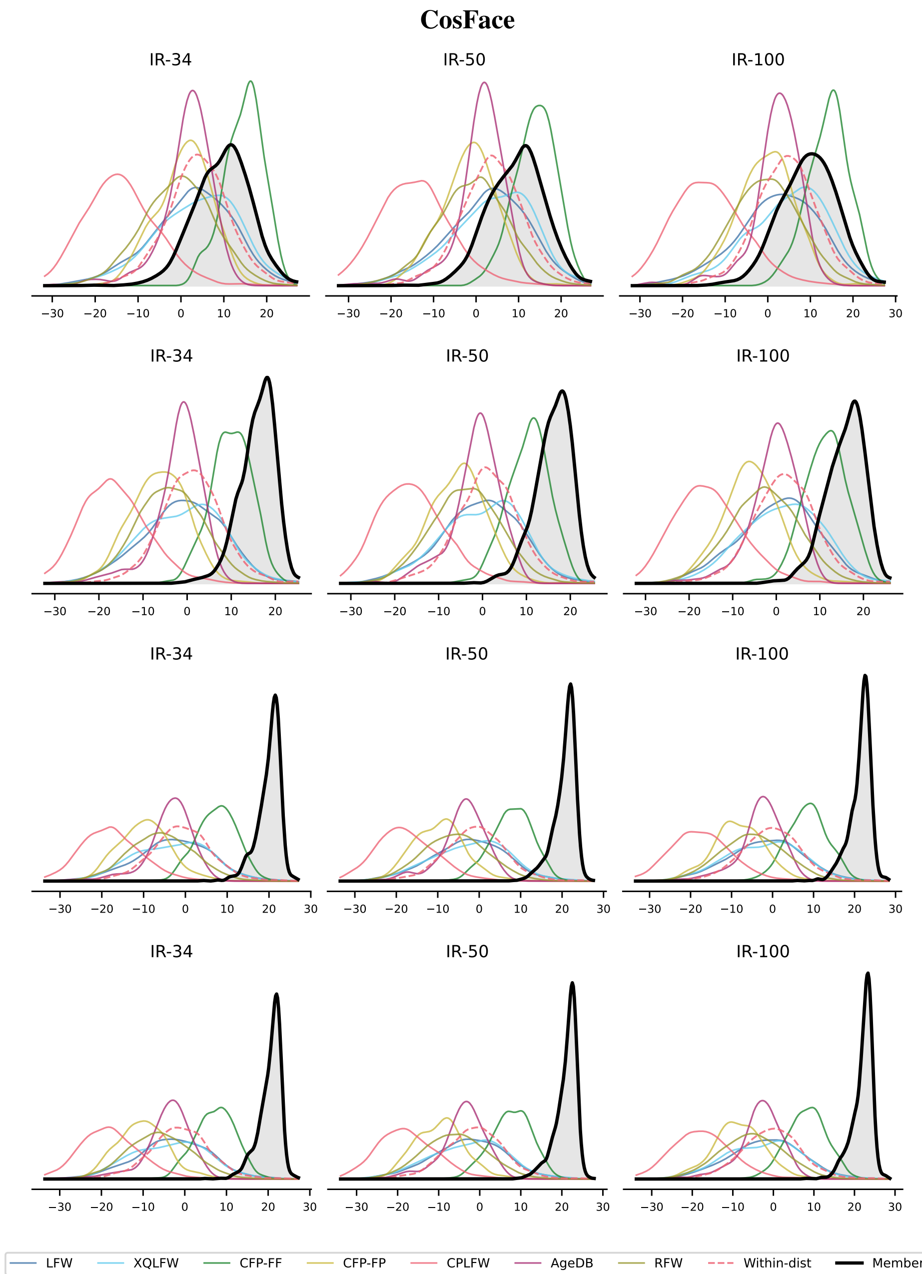


Figure 253. **Vary backbone — CosFace /** $n_{\text{ids}} = 1K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

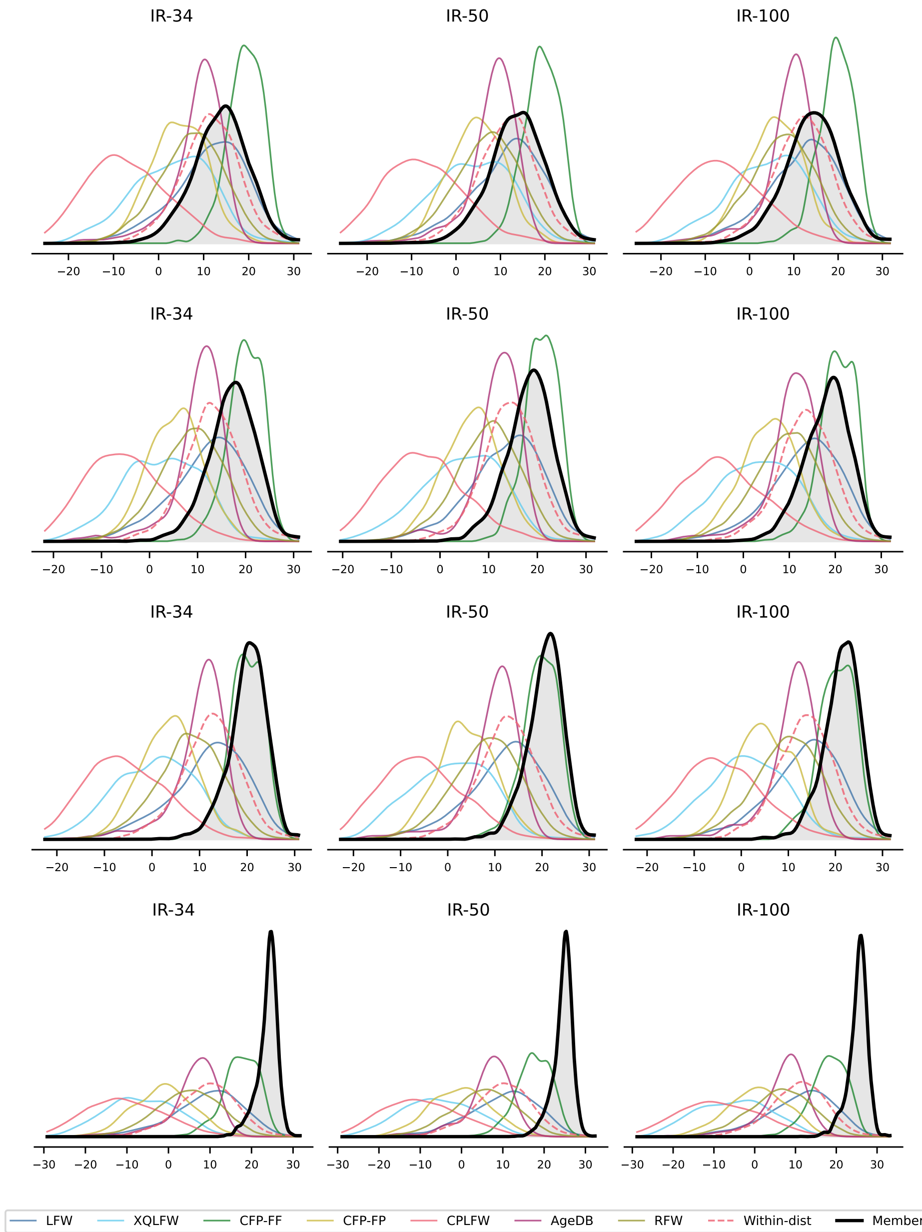


Figure 254. **Vary backbone — CosFace /** $n_{\text{ids}} = 10K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

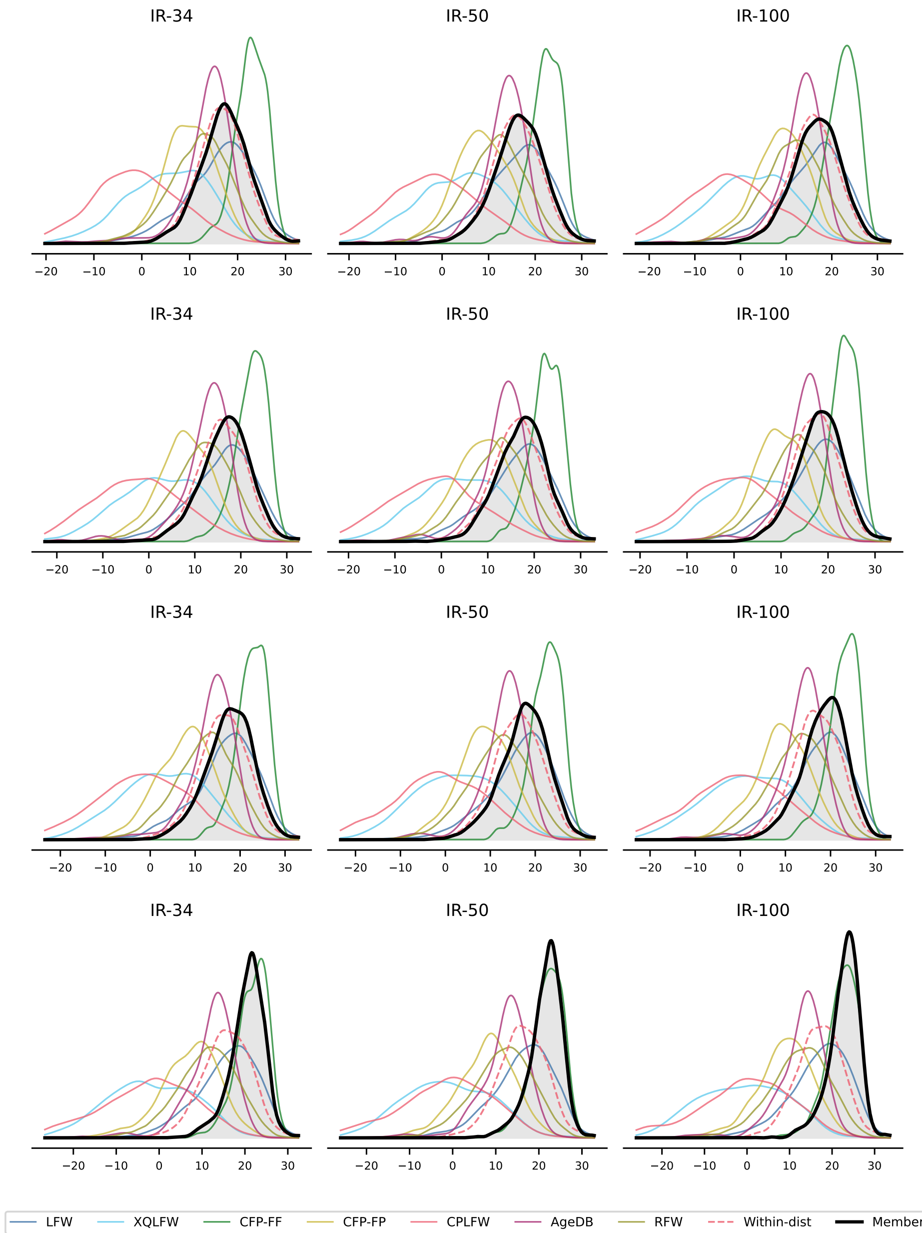


Figure 255. **Vary backbone — CosFace / $n_{\text{ids}} = 50K$.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

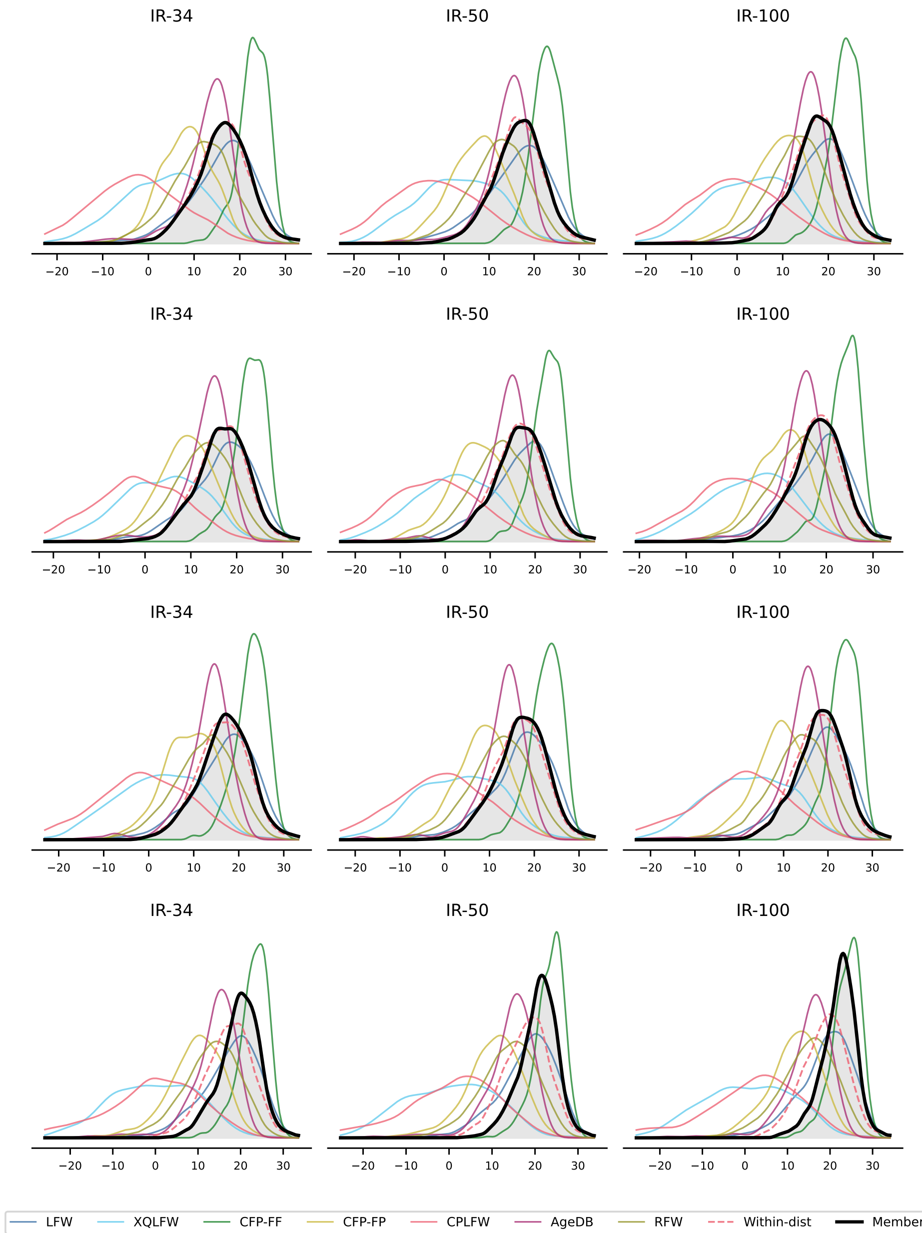


Figure 256. **Vary backbone — CosFace / $n_{\text{ids}} = 100K$.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

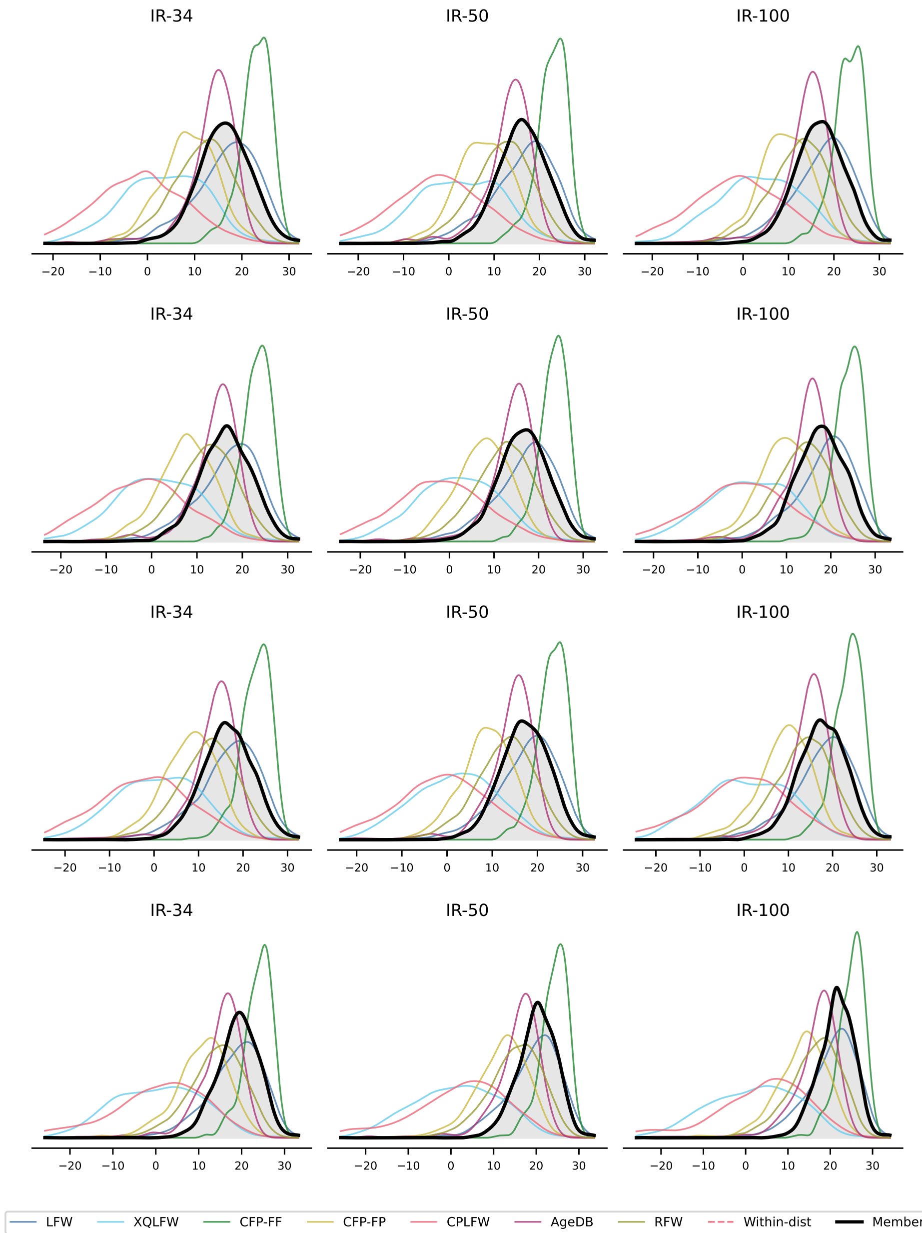


Figure 257. **Vary backbone — CosFace /** $n_{\text{ids}} = All$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

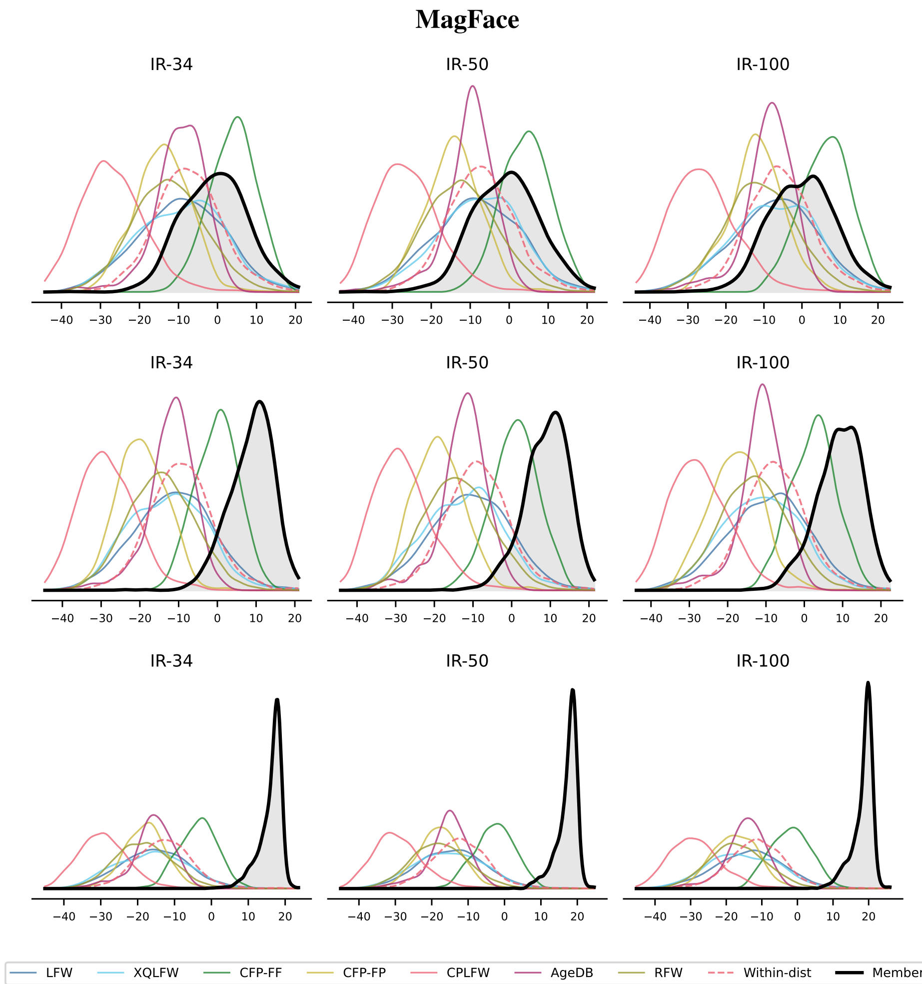


Figure 258. **Vary backbone — MagFace /** $n_{\text{ids}} = 1K$**.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

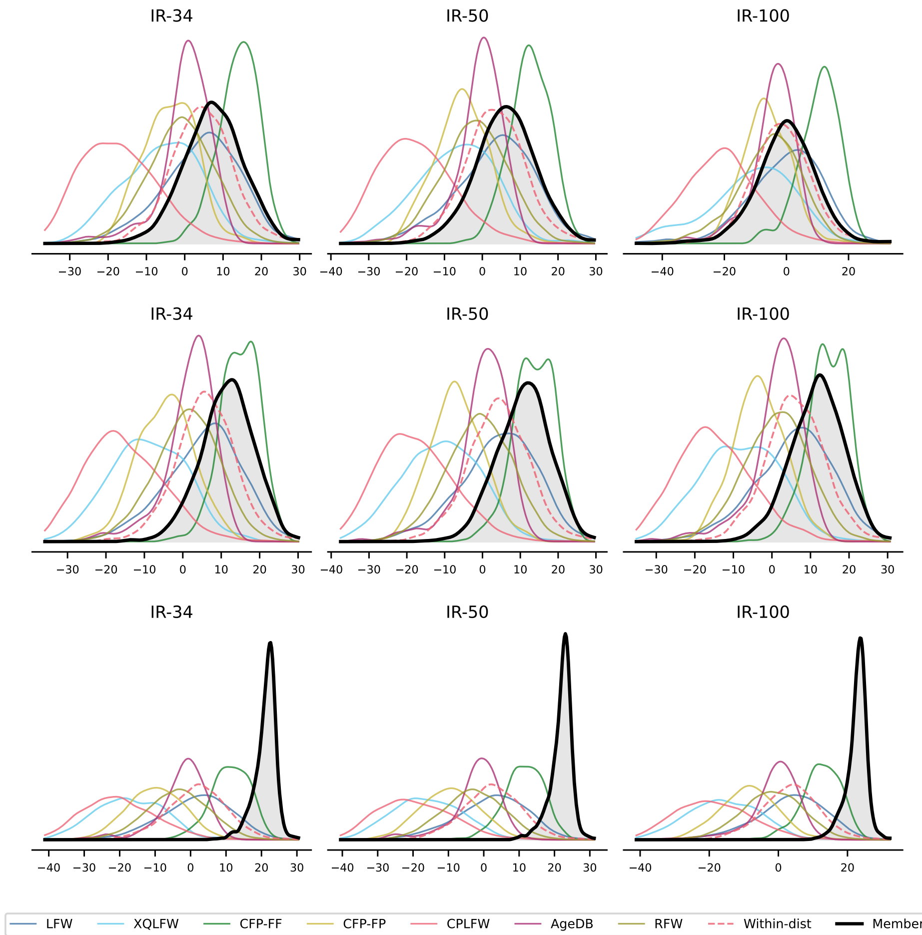


Figure 259. **Vary backbone — MagFace /** $n_{\mathbf{ids}} = 10K$**.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

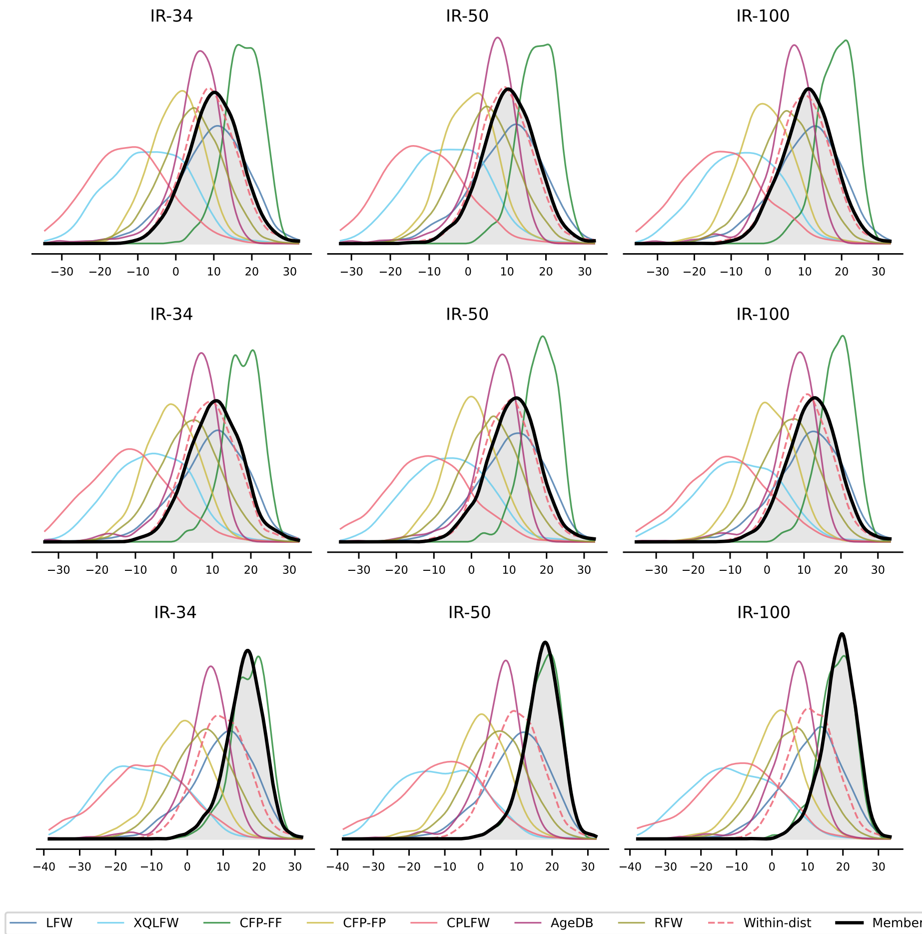


Figure 260. **Vary backbone — MagFace /** $n_{\text{ids}} = 50K$**.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

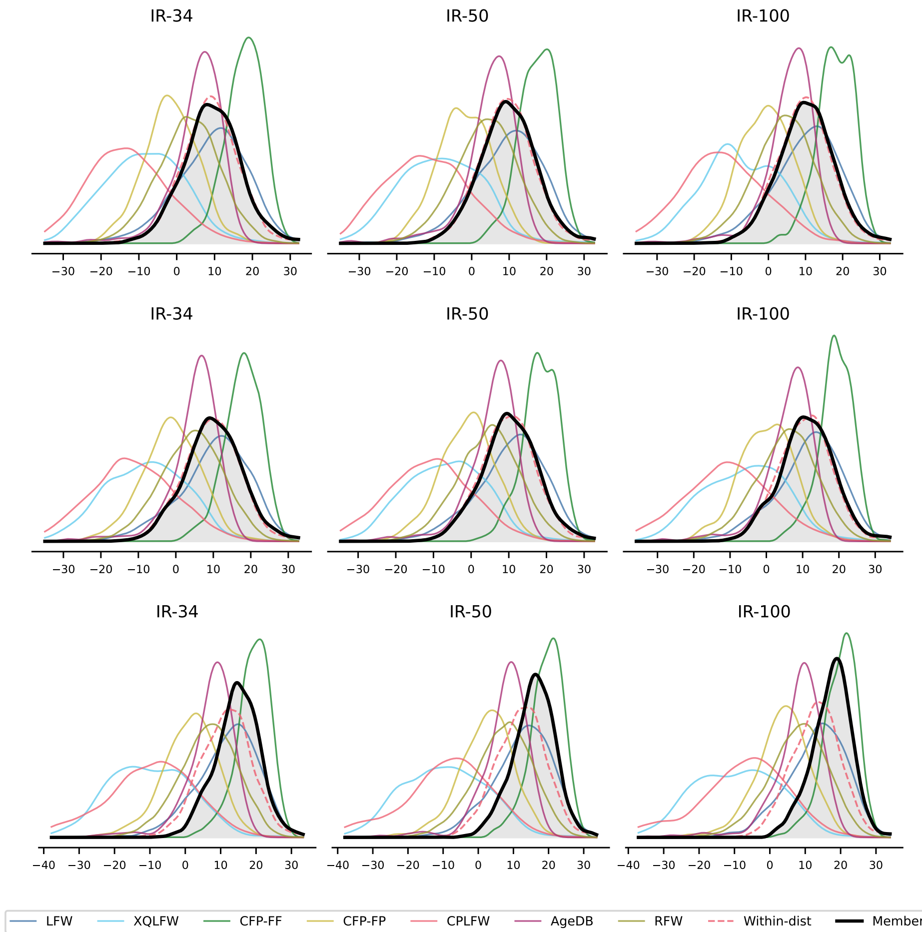


Figure 261. **Vary backbone — MagFace /** $n_{\text{ids}} = 100K$**.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

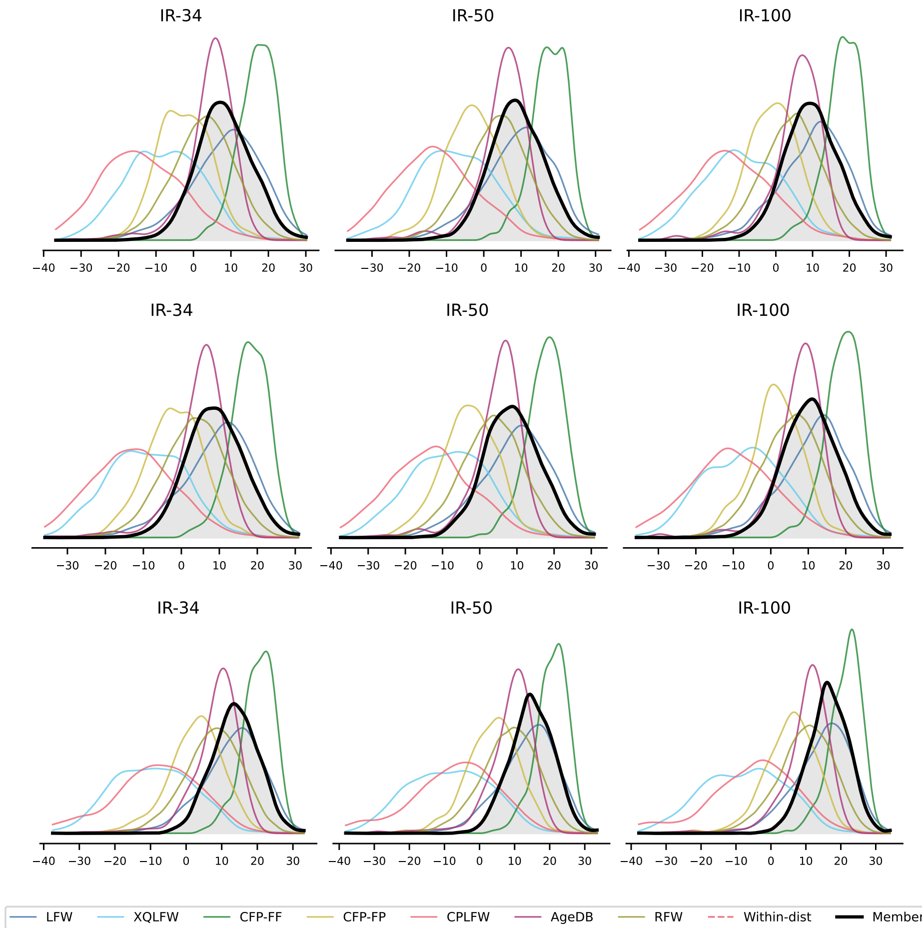


Figure 262. **Vary backbone — MagFace /** $n_{\textbf{ids}} = All$**.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

7. Proto-softmax CE distributions

7.1. Vary training duration

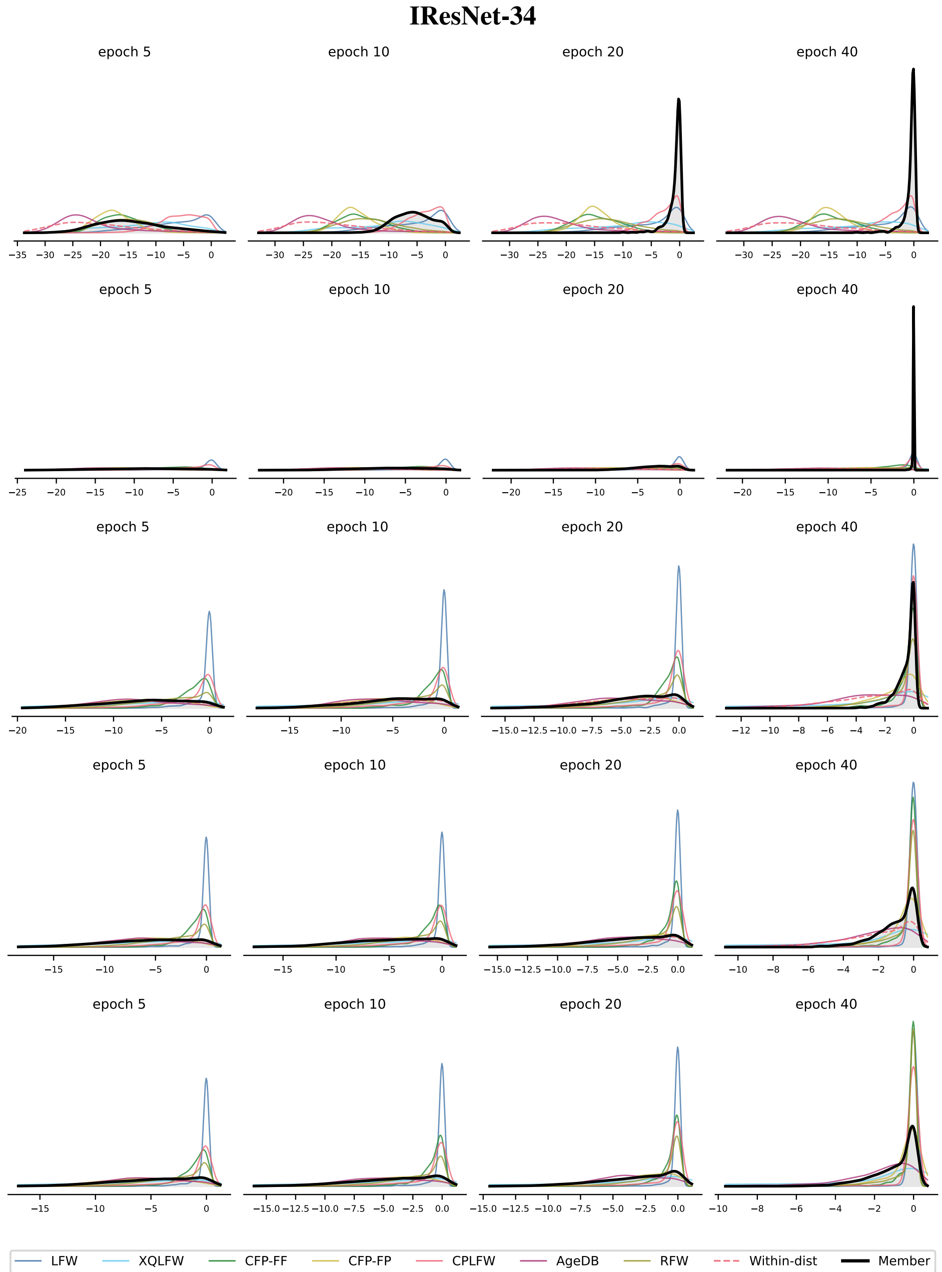


Figure 263. **Vary training duration — IResNet-34 / ArcFace.** Rows: $n_{\text{ids}} \in \{1\text{K},\ 10\text{K},\ 50\text{K},\ 100\text{K},\ \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

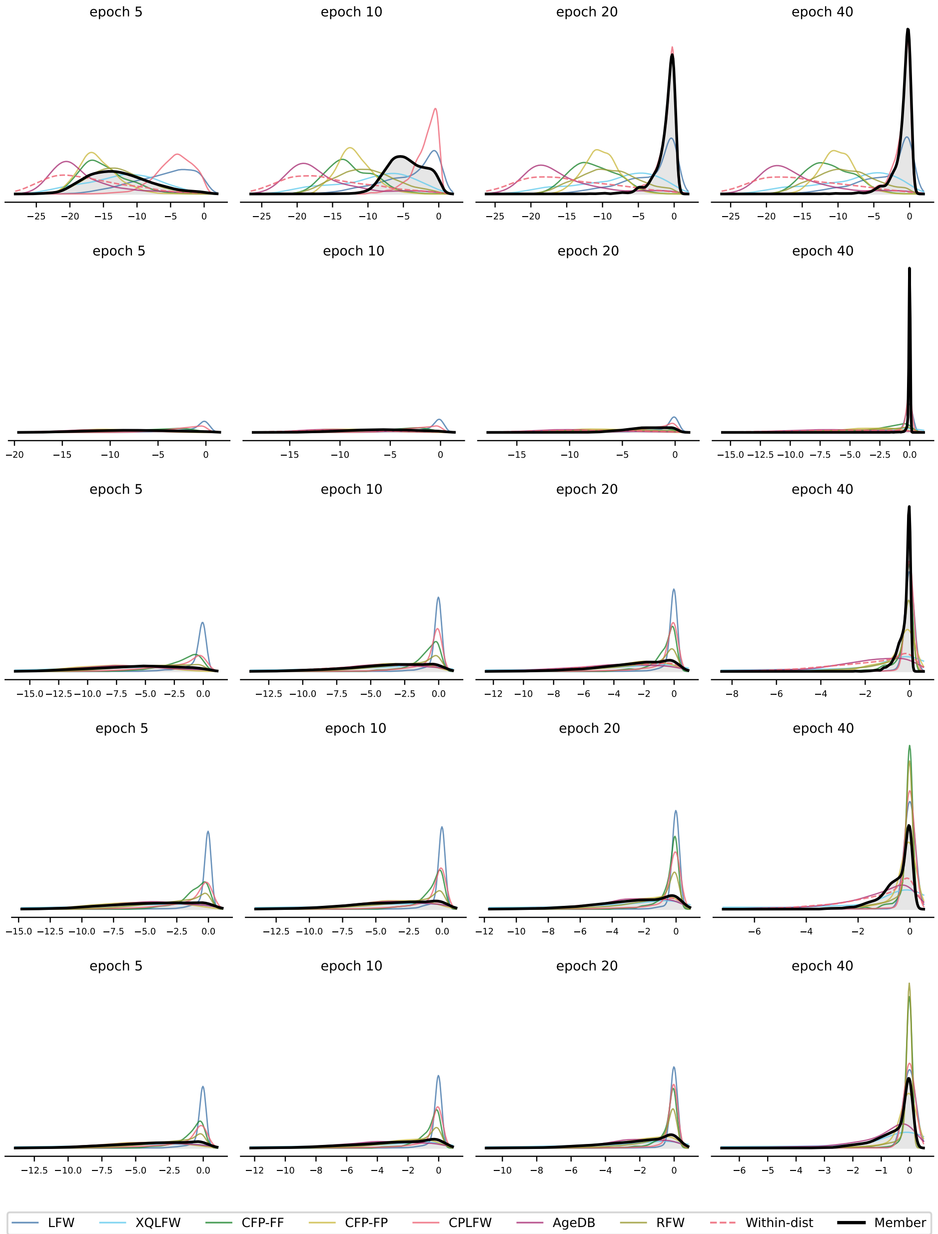


Figure 264. **Vary training duration — IResNet-34 / CosFace.** Rows: $n_{\text{ids}} \in \{1\text{K},\ 10\text{K},\ 50\text{K},\ 100\text{K},\ \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

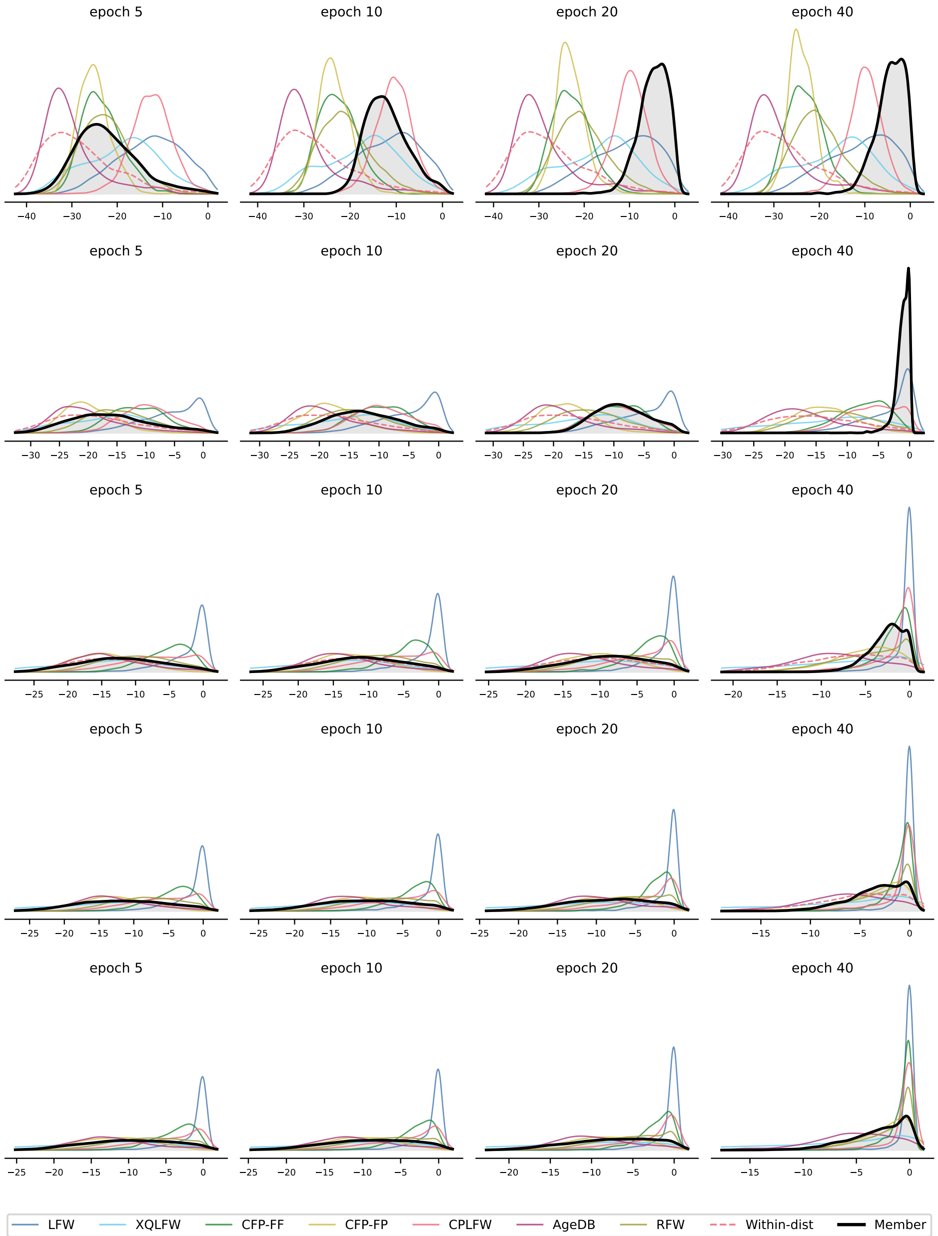


Figure 265. **Vary training duration — IResNet-34 / MagFace.** Rows: $n_{\text{ids}} \in \{1\text{K},\ 10\text{K},\ 50\text{K},\ 100\text{K},\ \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

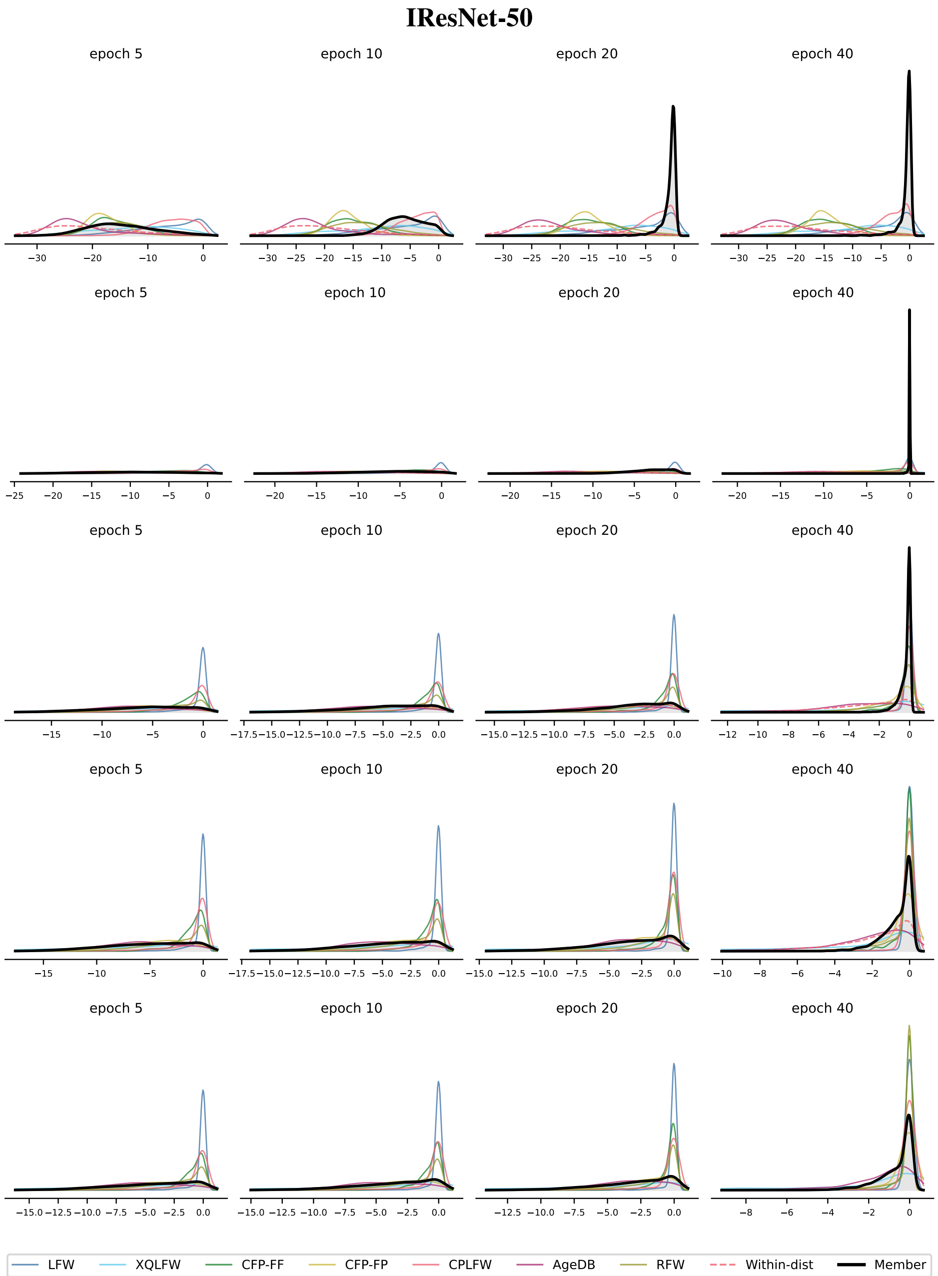


Figure 266. **Vary training duration — IResNet-50 / ArcFace.** Rows: $n_{\text{ids}} \in \{1\text{K},\ 10\text{K},\ 50\text{K},\ 100\text{K},\ \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

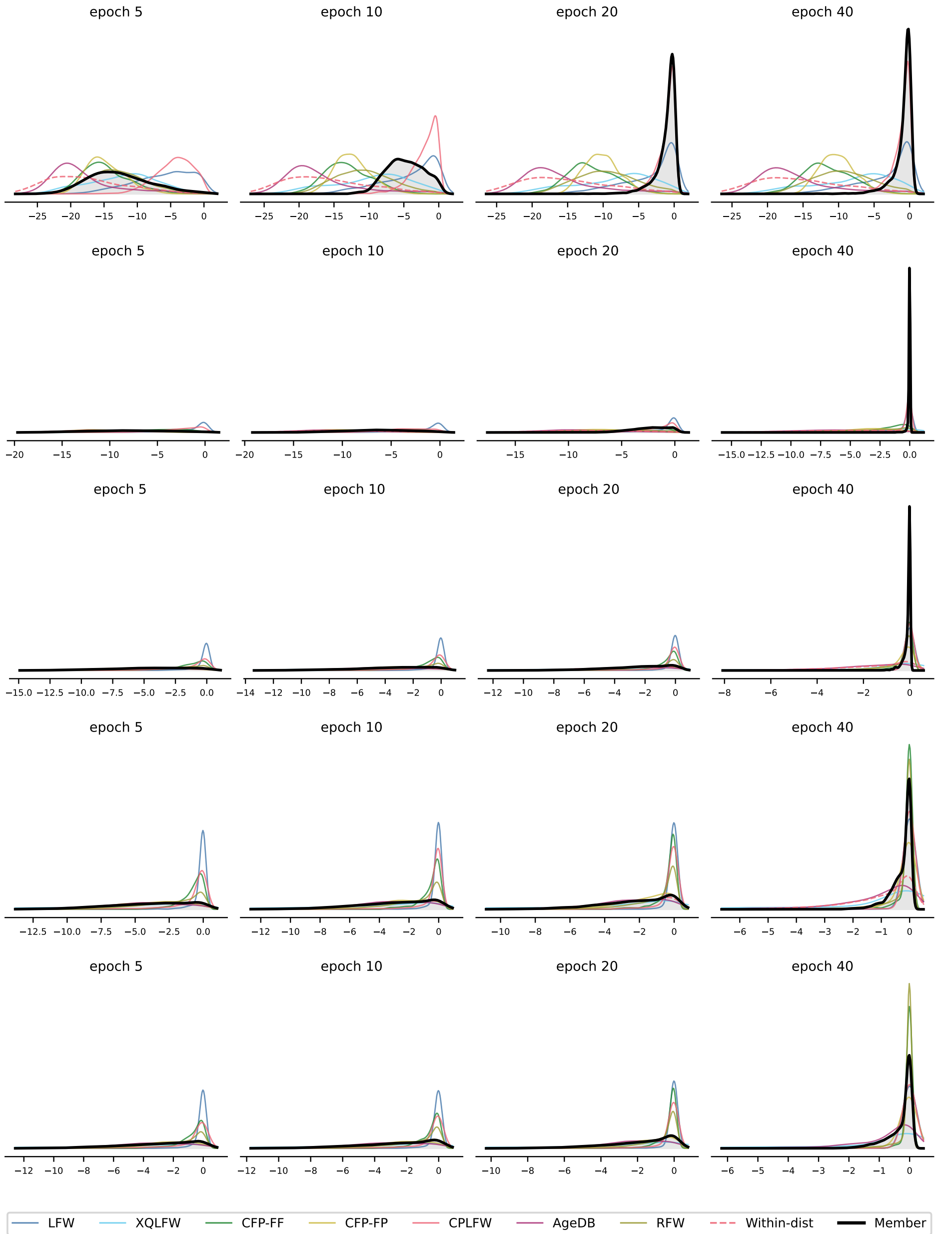


Figure 267. **Vary training duration — IResNet-50 / CosFace.** Rows: $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

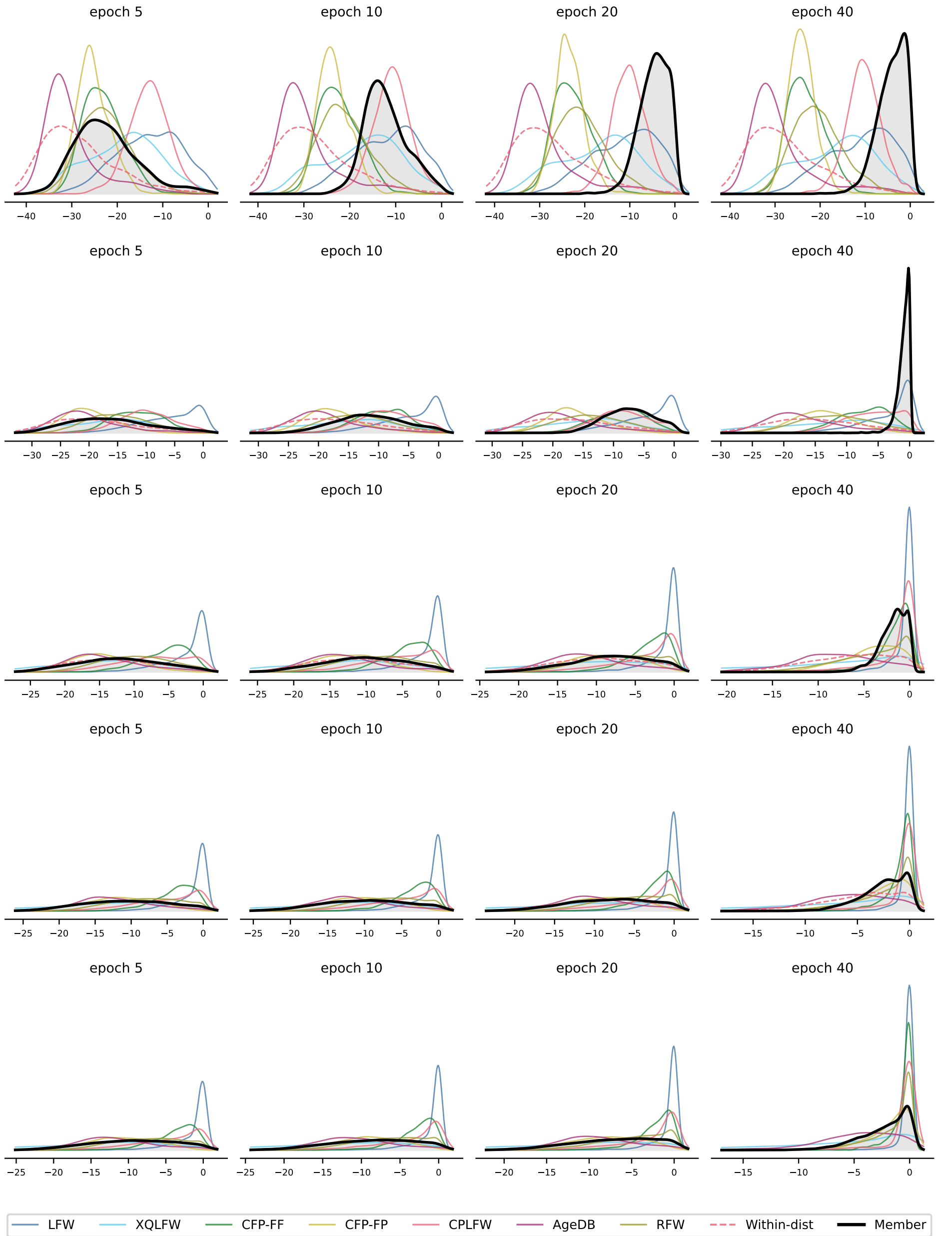


Figure 268. **Vary training duration — IResNet-50 / MagFace.** Rows: $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

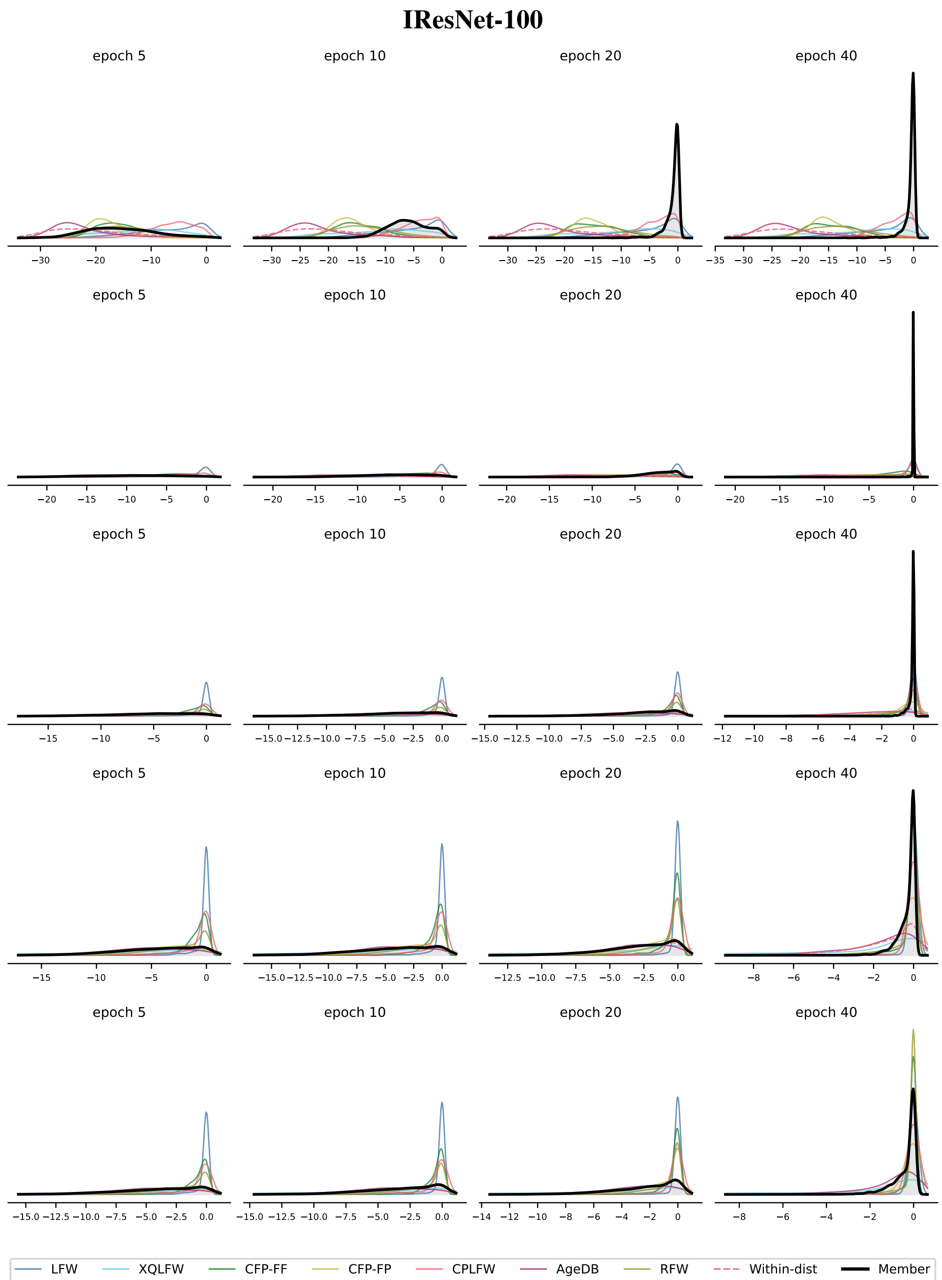


Figure 269. **Vary training duration — IResNet-100 / ArcFace.** Rows: $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

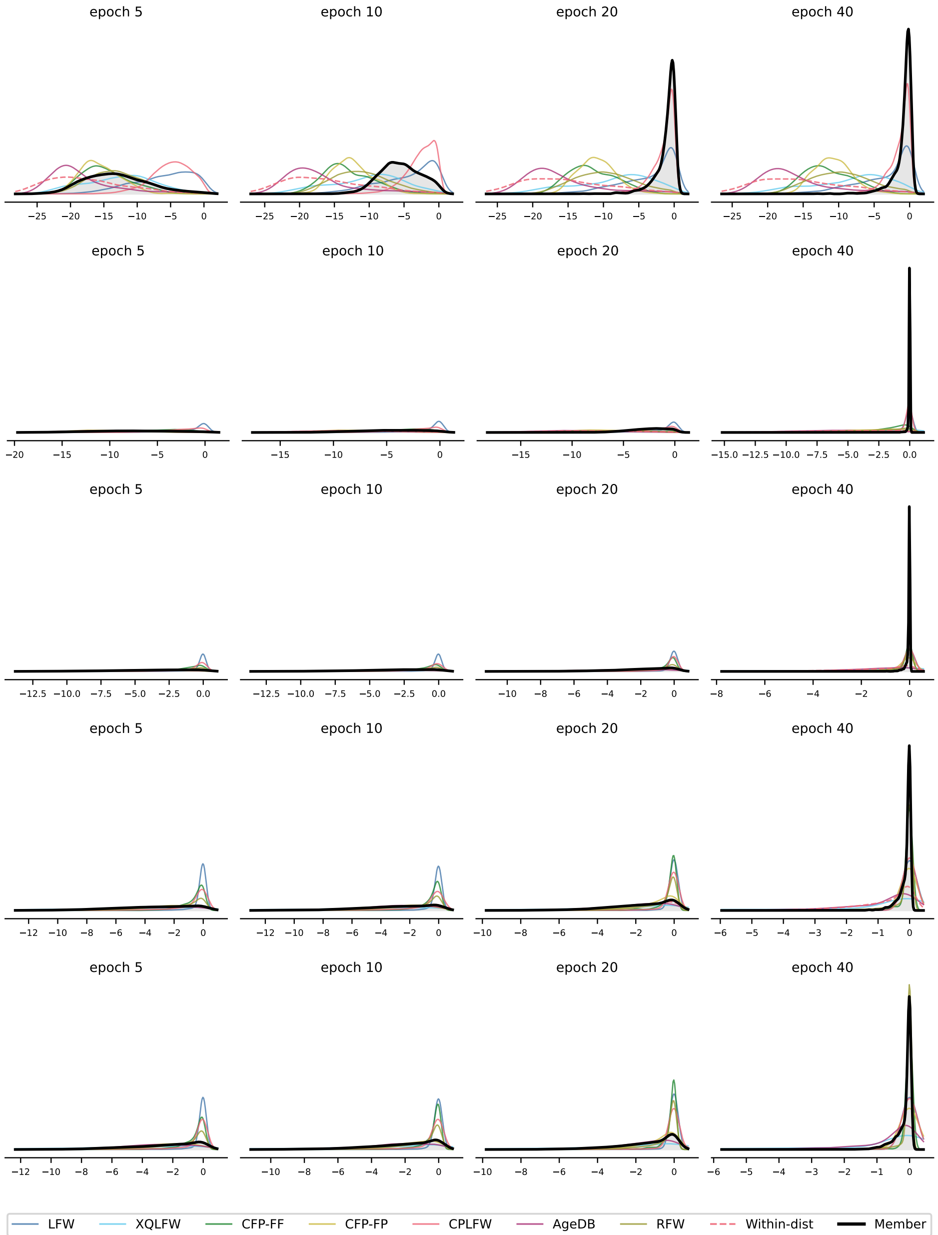


Figure 270. **Vary training duration — IResNet-100 / CosFace.** Rows: $n_{\text{ids}} \in \{1\text{K},\ 10\text{K},\ 50\text{K},\ 100\text{K},\ \text{All}\}$ (top to bottom). Within each row, epochs 5, 10, 20, and 40 (left to right).

7.2. Vary training-set size

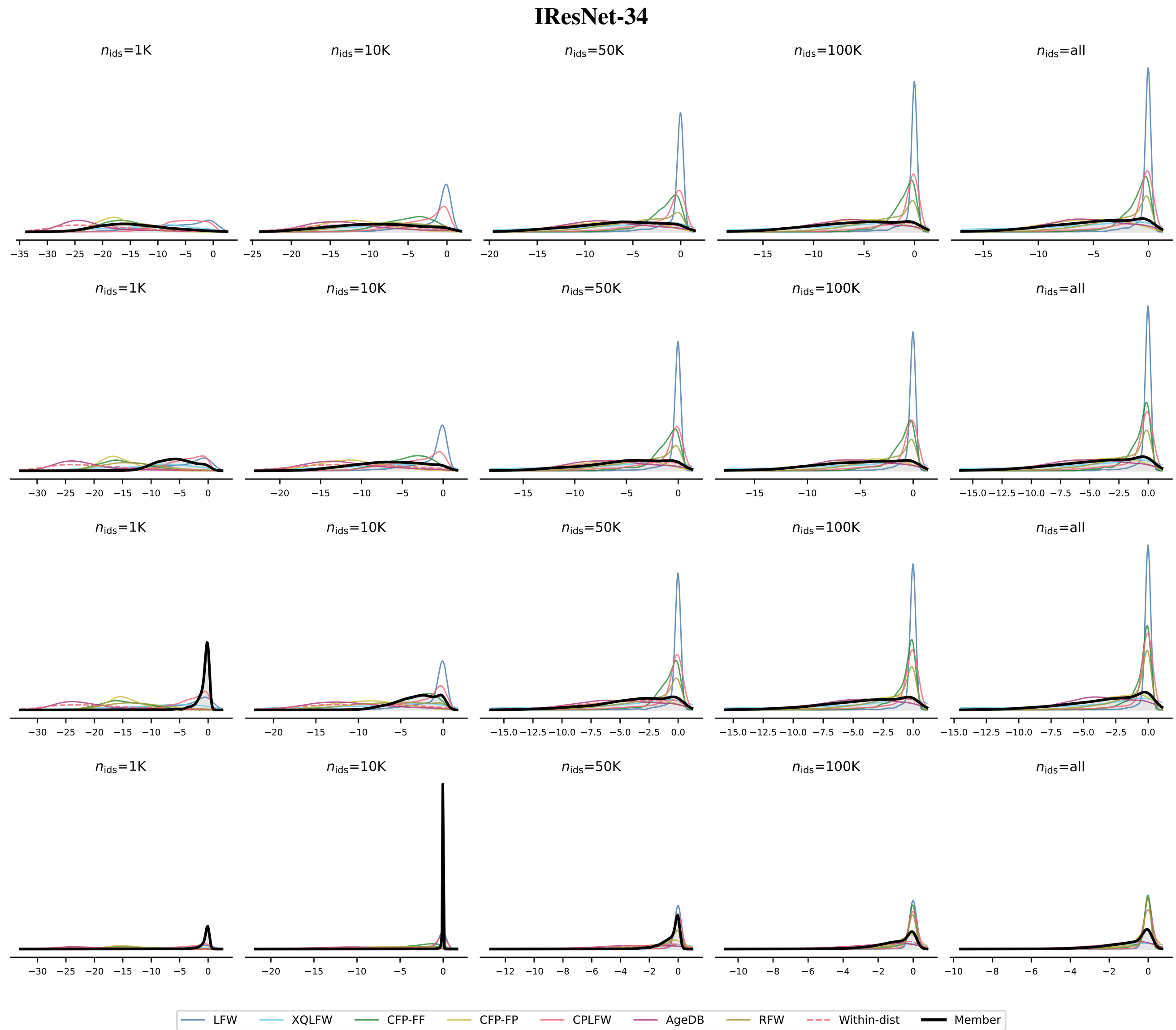


Figure 271. **Vary training-set size — IResNet-34 / ArcFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

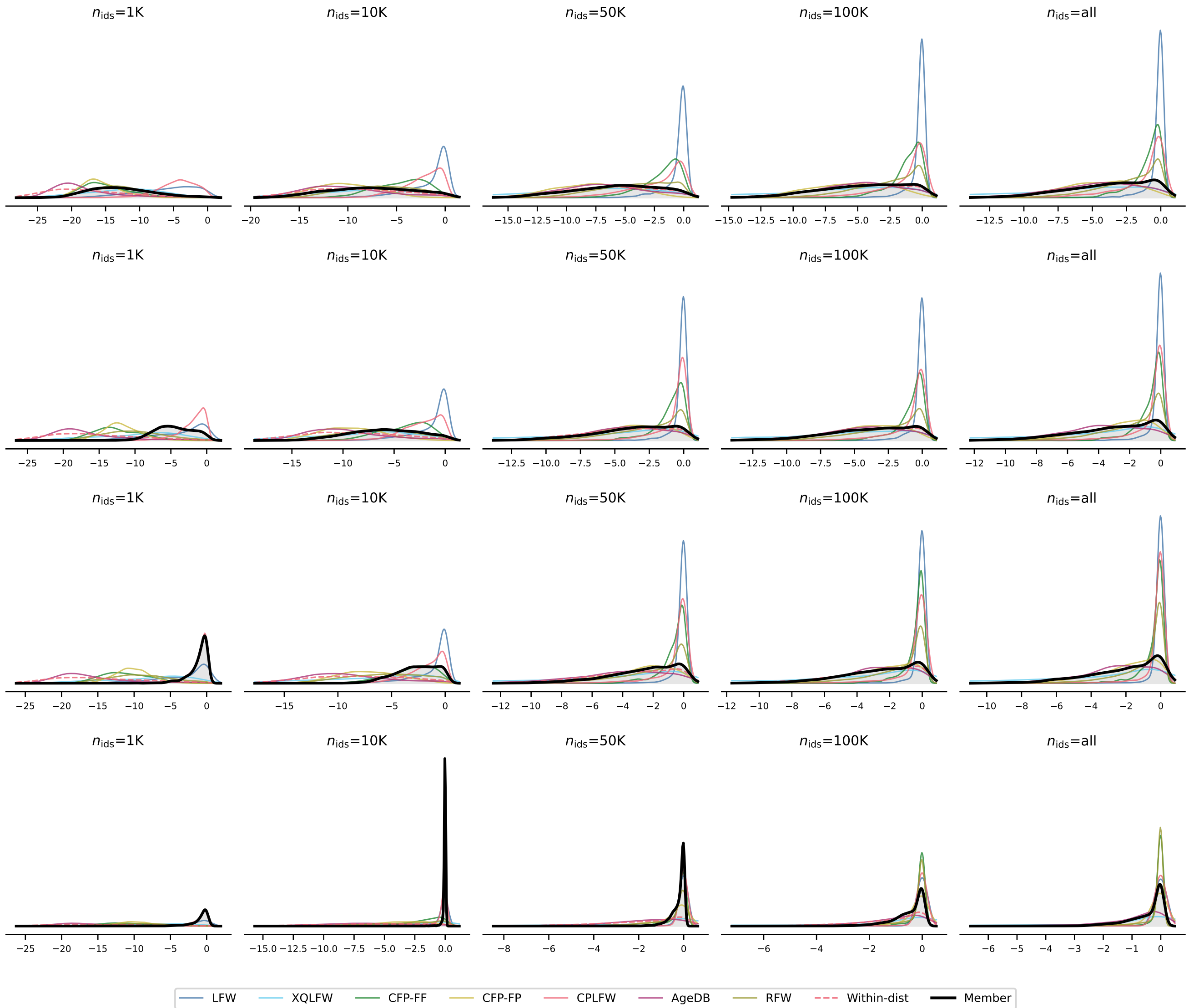


Figure 272. **Vary training-set size — IResNet-34 / CosFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

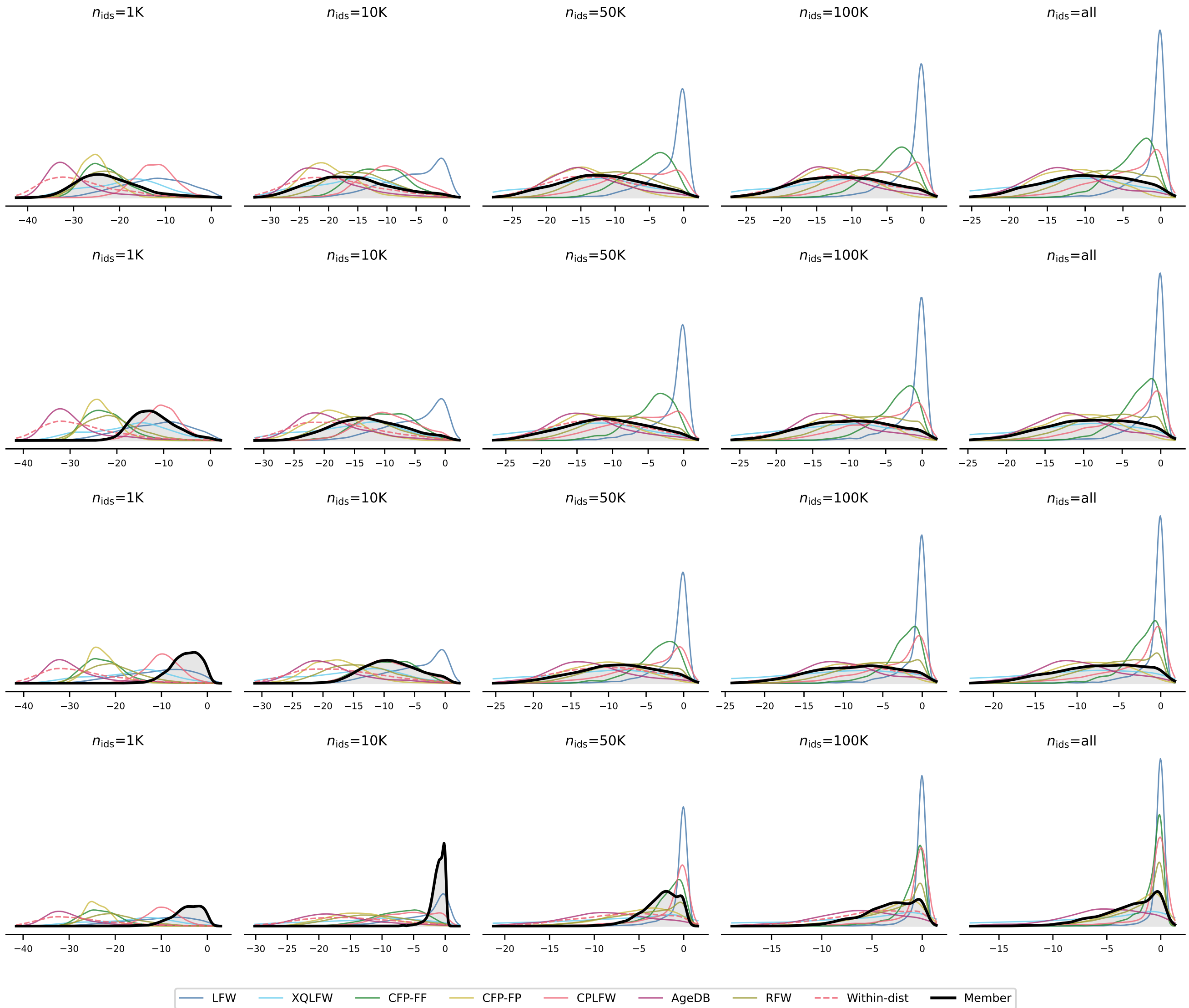


Figure 273. **Vary training-set size — IResNet-34 / MagFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

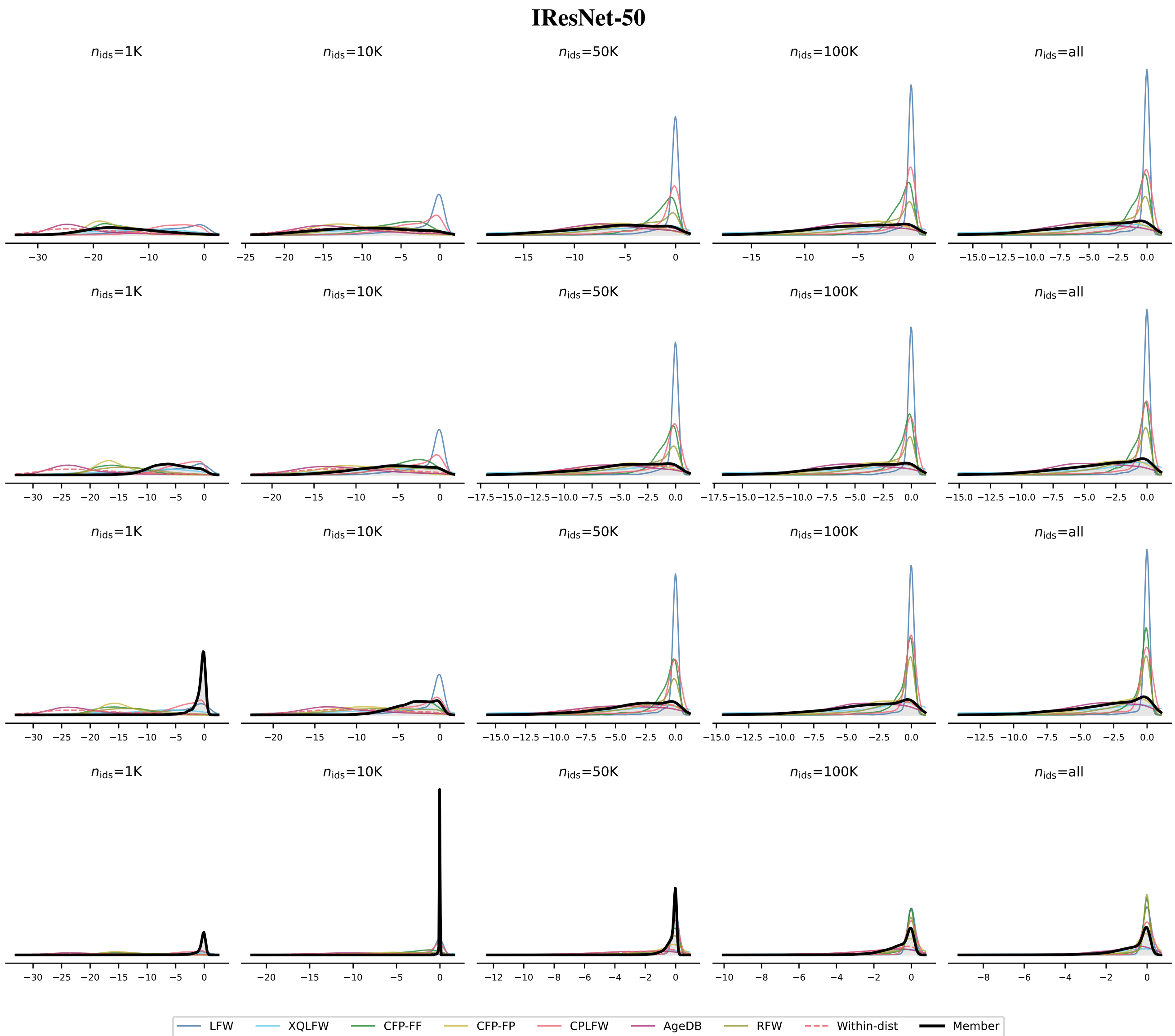


Figure 274. **Vary training-set size — IResNet-50 / ArcFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

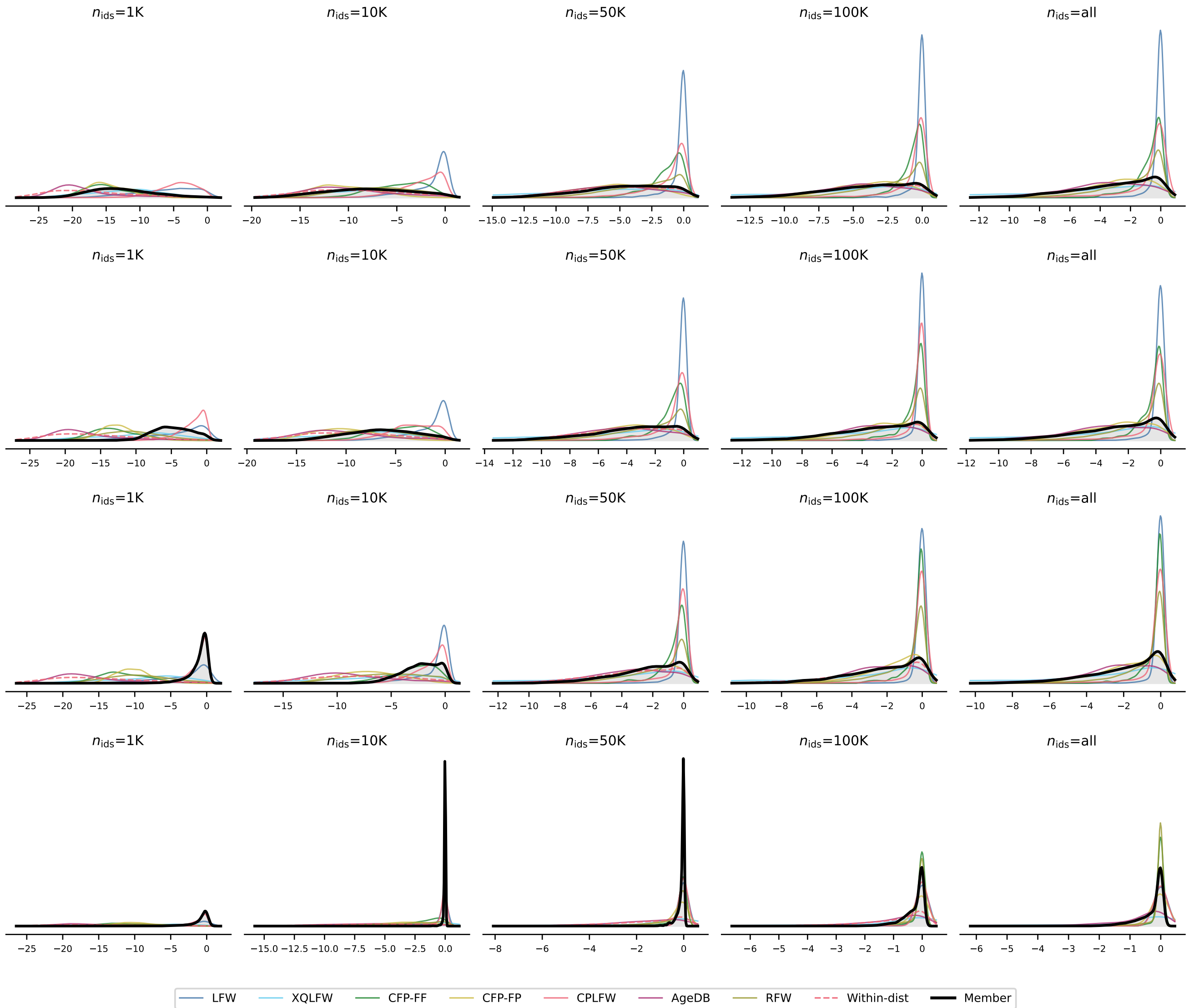


Figure 275. **Vary training-set size — IResNet-50 / CosFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

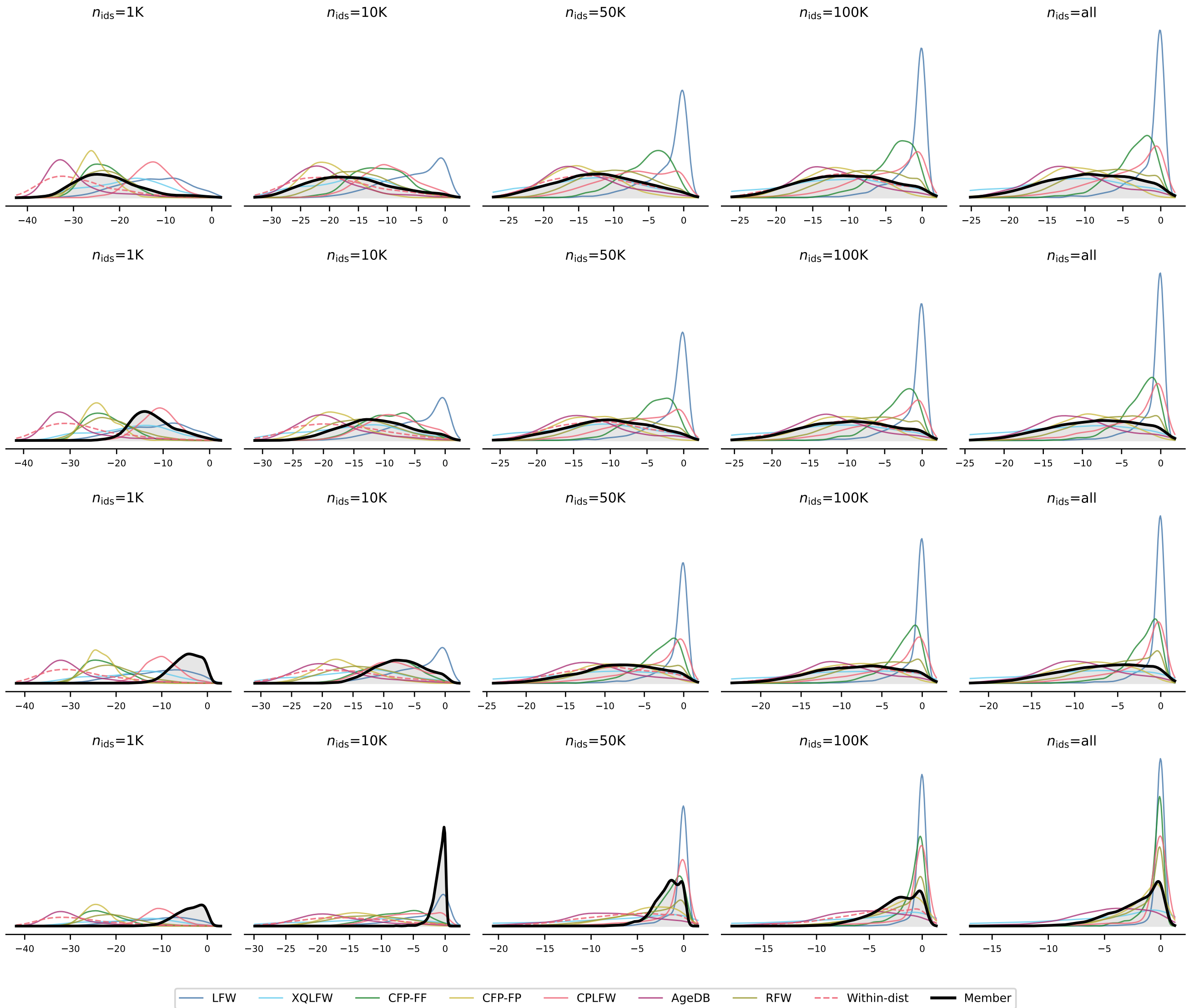


Figure 276. **Vary training-set size — IResNet-50 / MagFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

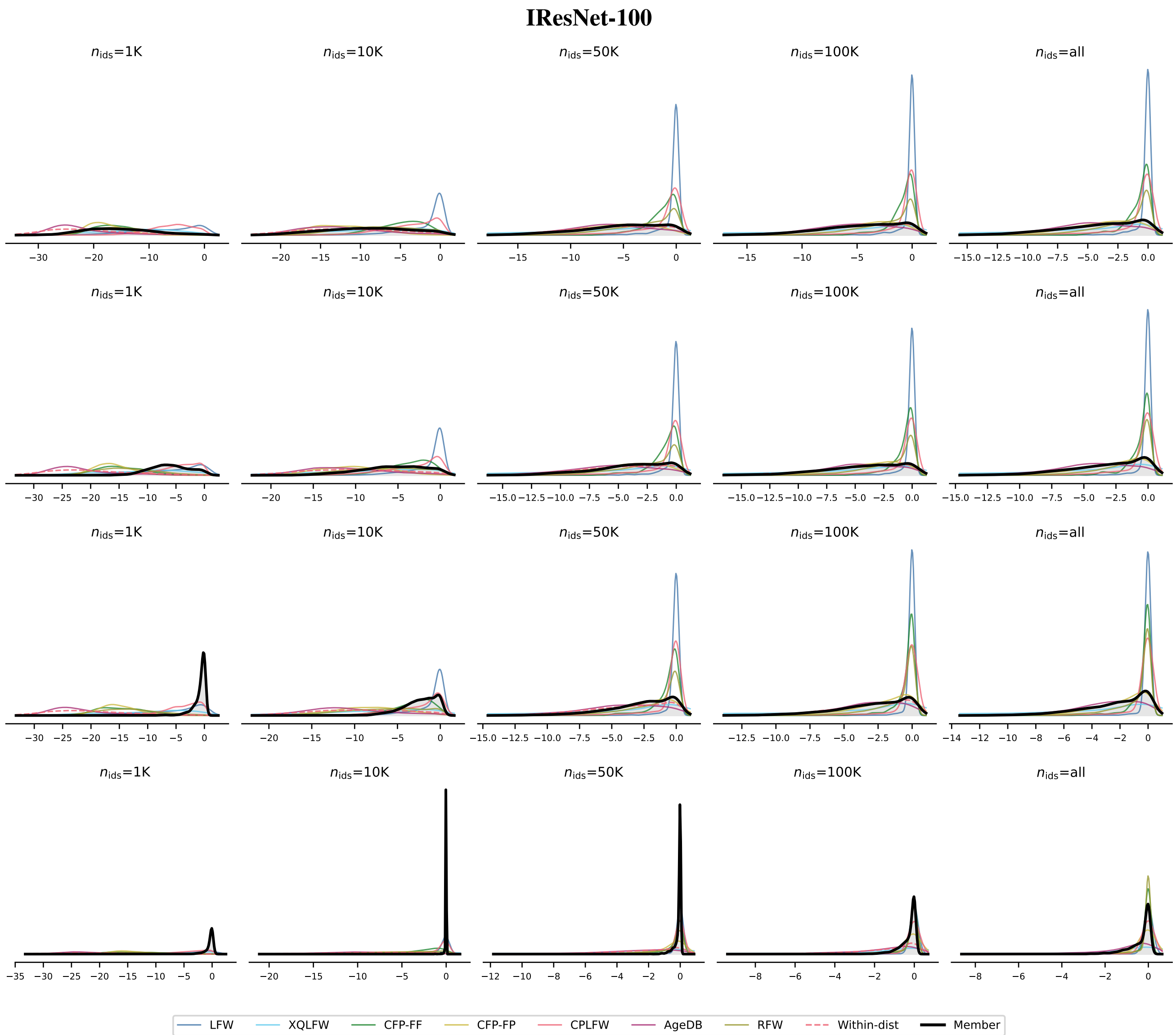


Figure 277. **Vary training-set size — IResNet-100 / ArcFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

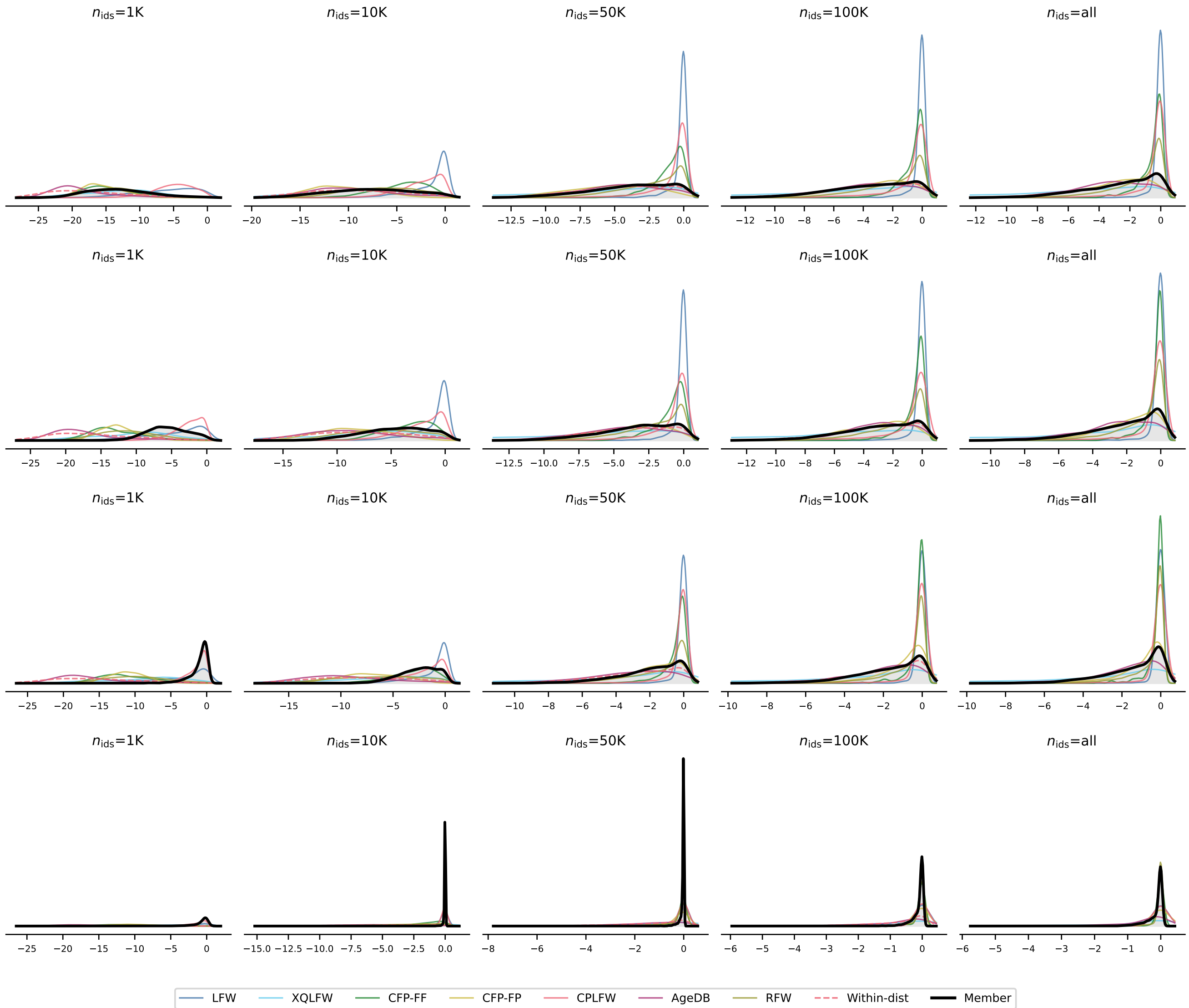


Figure 278. **Vary training-set size — IResNet-100 / CosFace.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

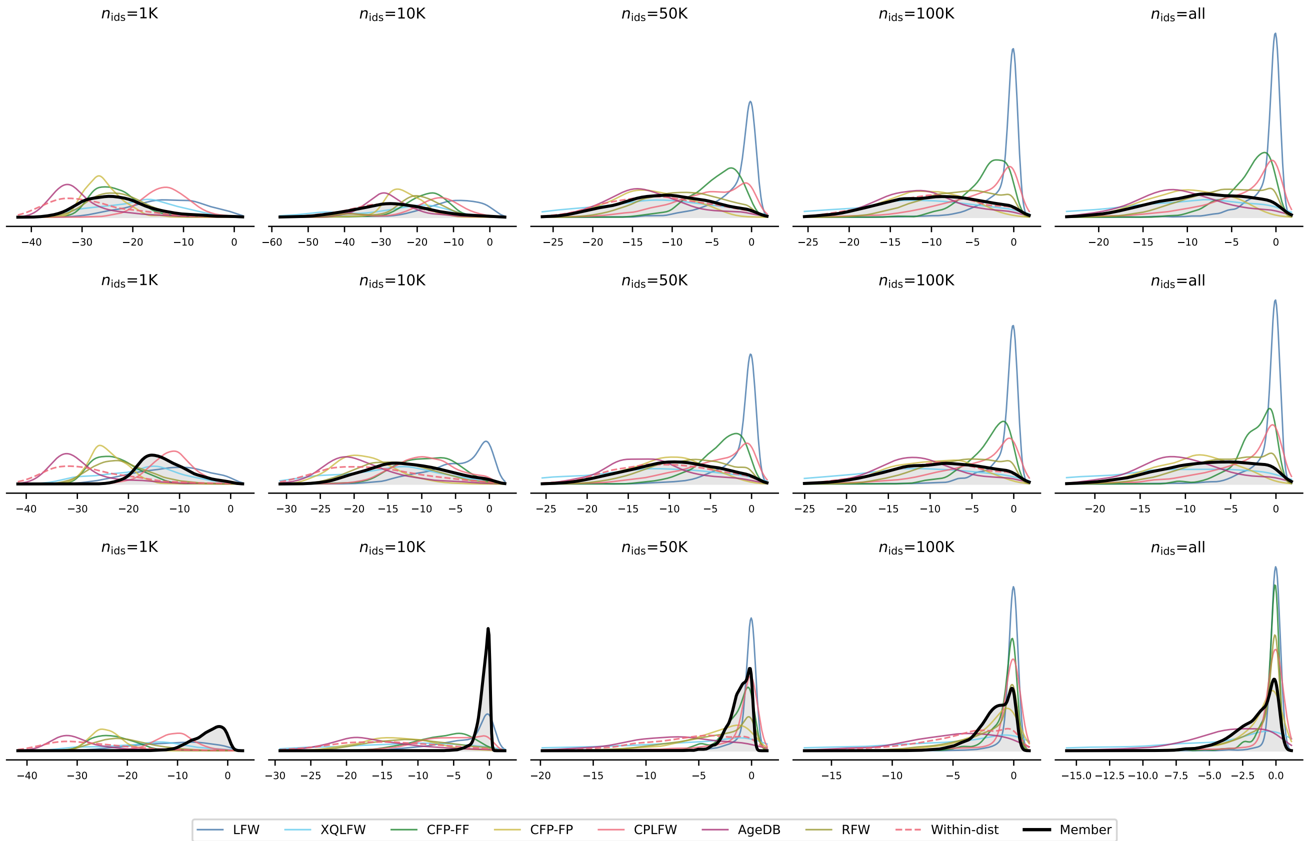


Figure 279. **Vary training-set size — IResNet-100 / MagFace.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, $n_{\text{ids}} \in \{1\text{K}, 10\text{K}, 50\text{K}, 100\text{K}, \text{All}\}$ (left to right).

7.3. Vary loss head

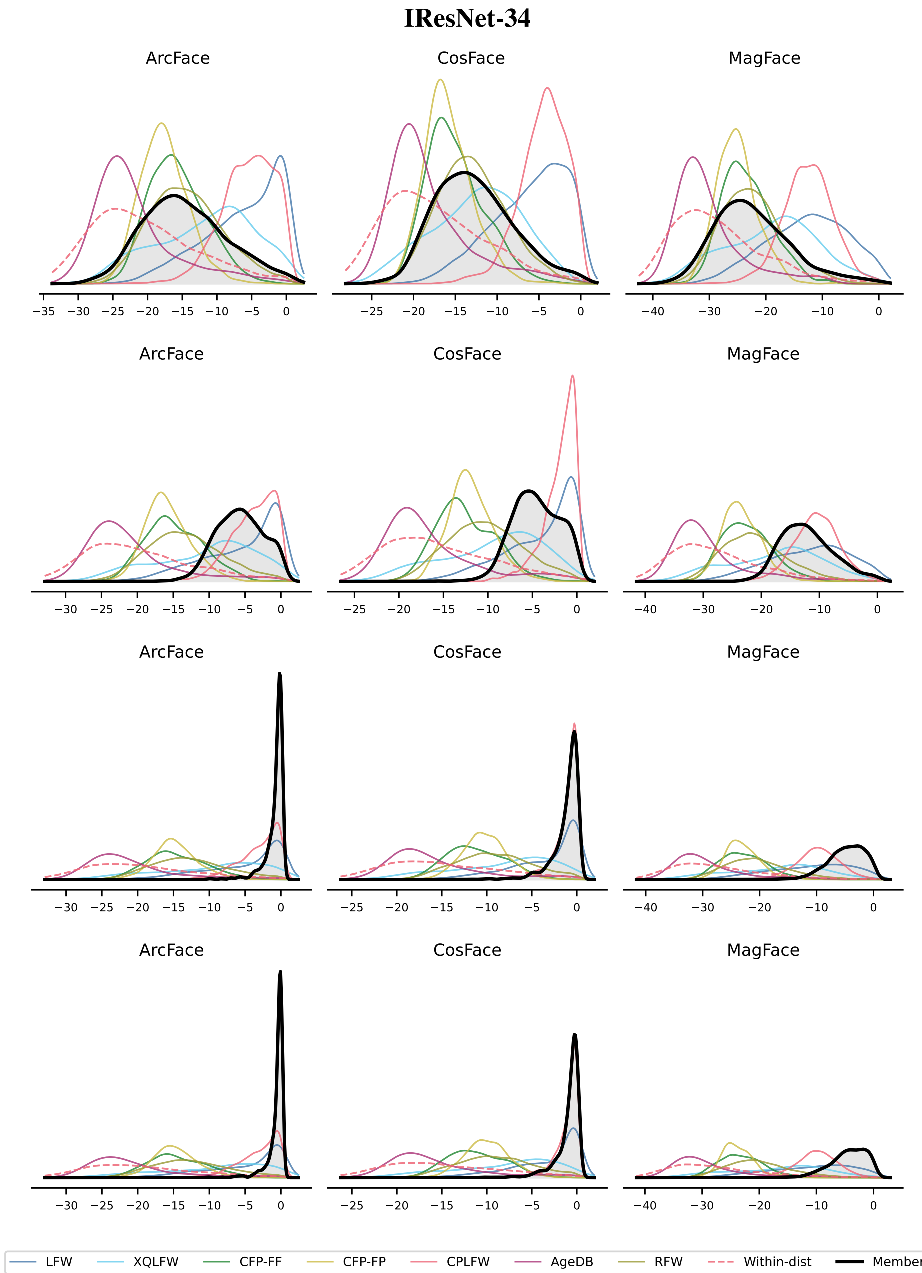


Figure 280. **Vary loss head — IResNet-34 /** $n_{\text{ids}} = 1K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

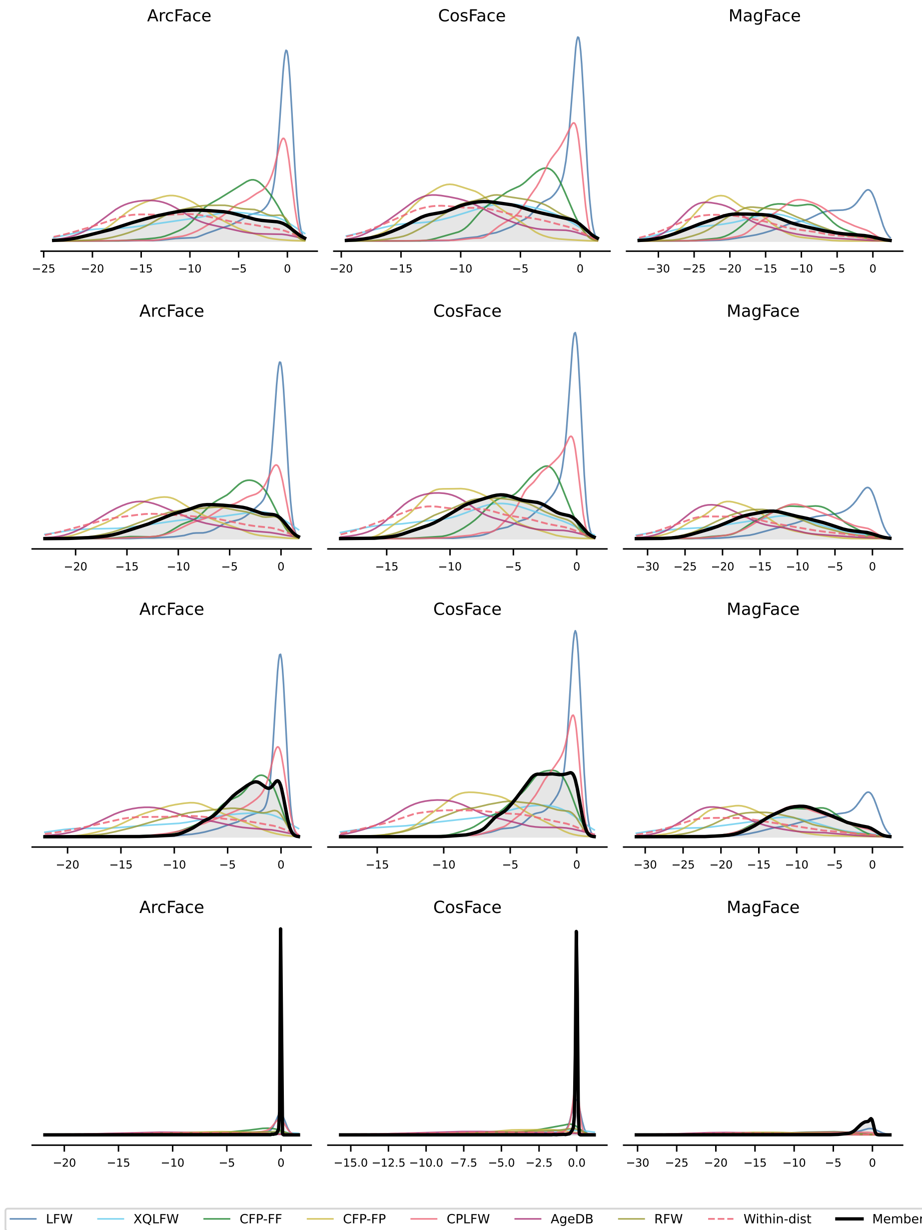


Figure 281. **Vary loss head — IResNet-34 /** $n_{\text{ids}} = 10K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

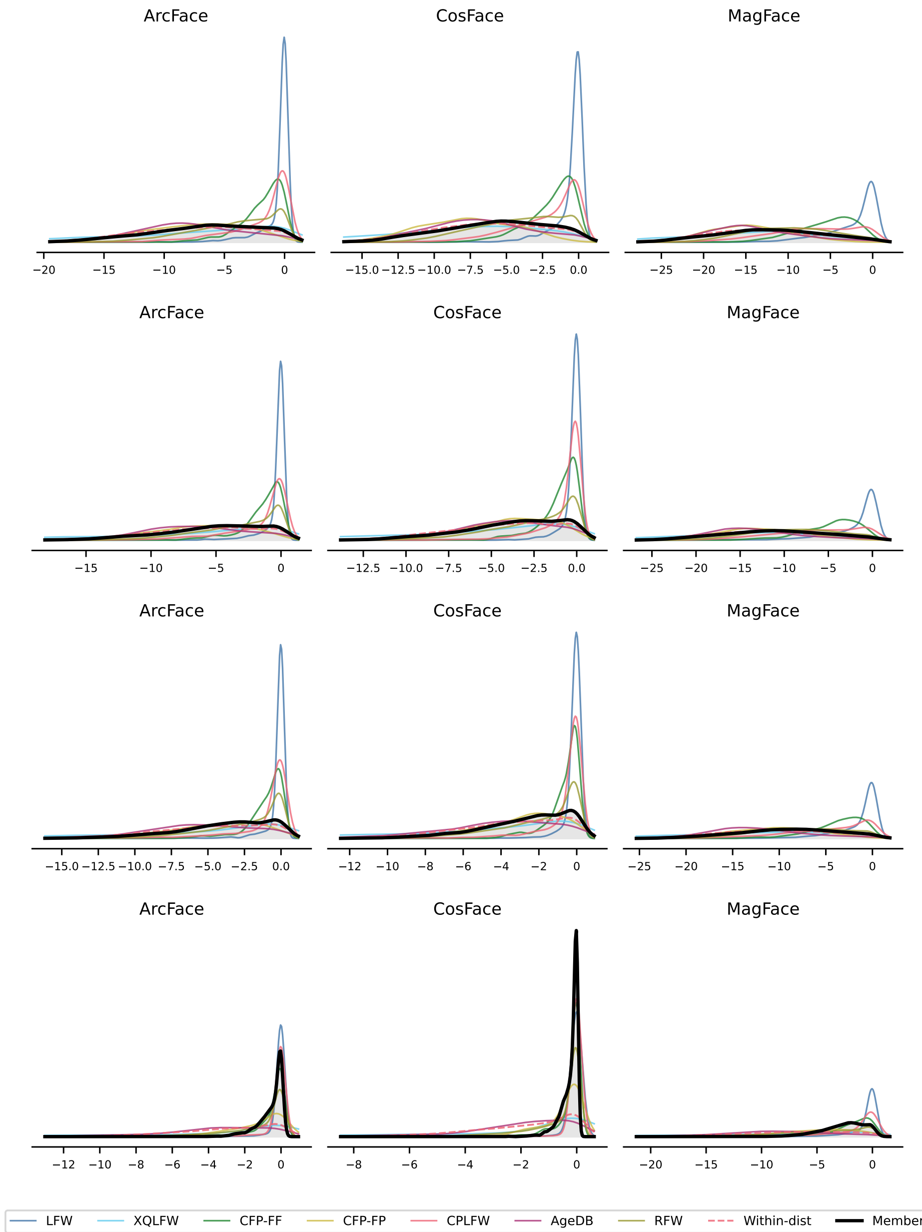


Figure 282. **Vary loss head — IResNet-34 / $n_{\text{ids}} = 50K$.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

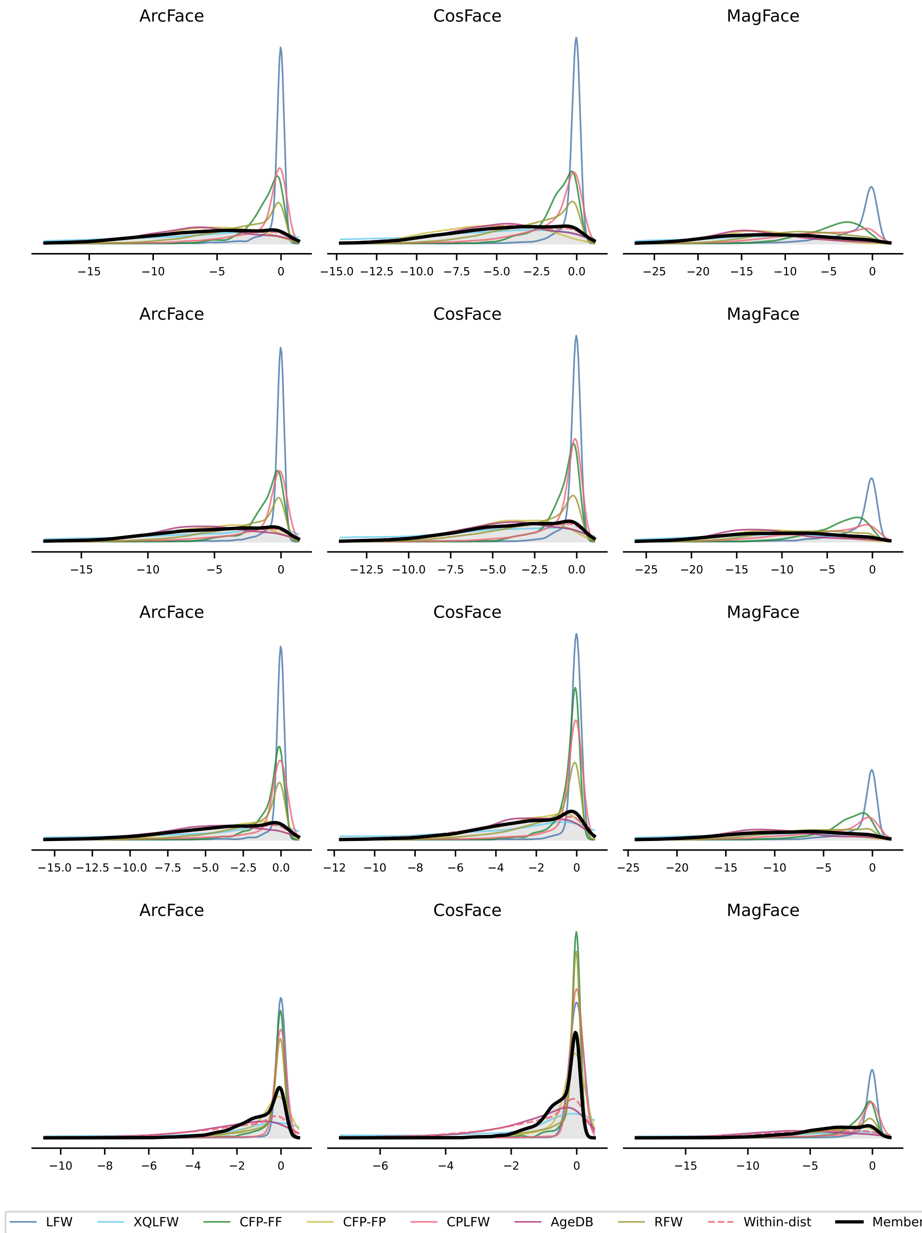


Figure 283. **Vary loss head — IResNet-34 /** $n_{\text{ids}} = 100K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

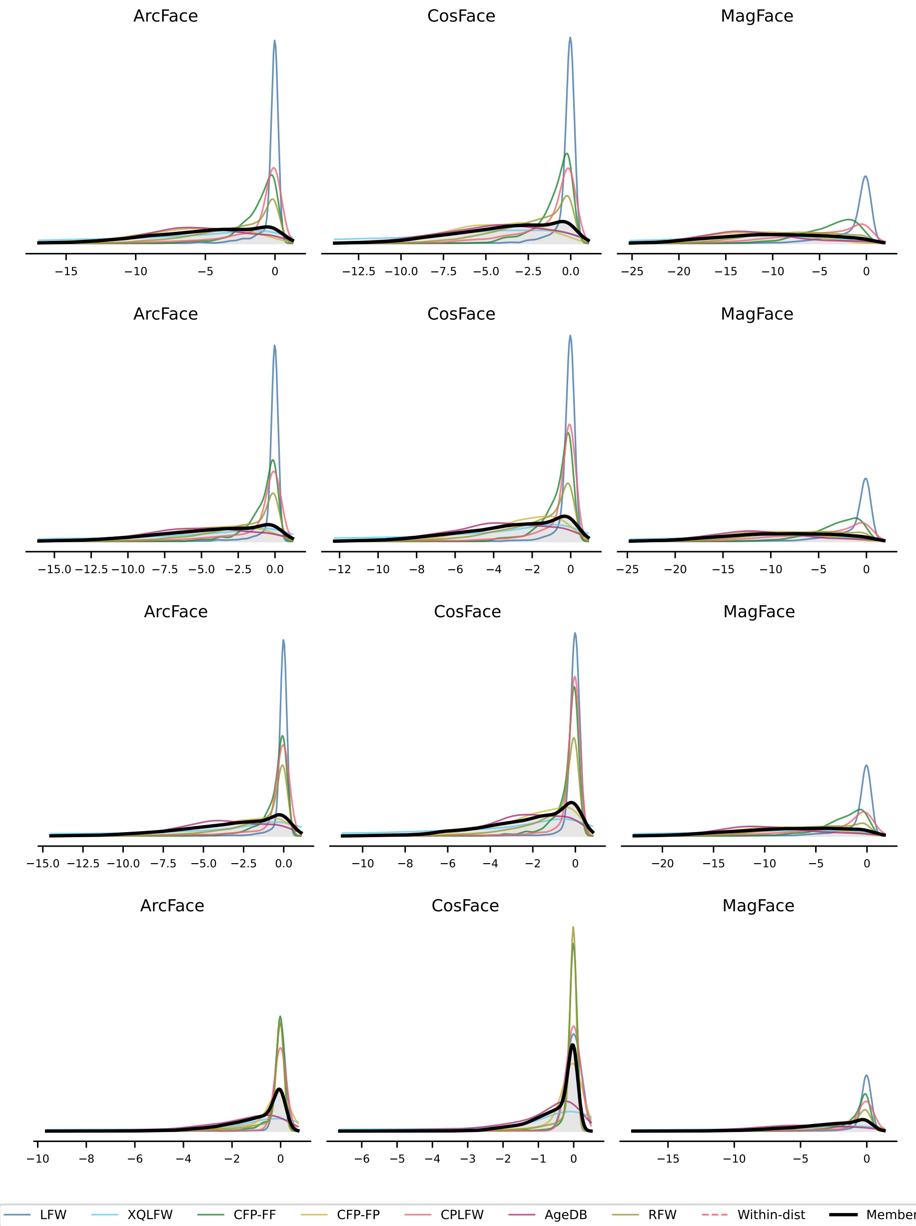


Figure 284. **Vary loss head — IResNet-34 /** $n_{\text{ids}} = All$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

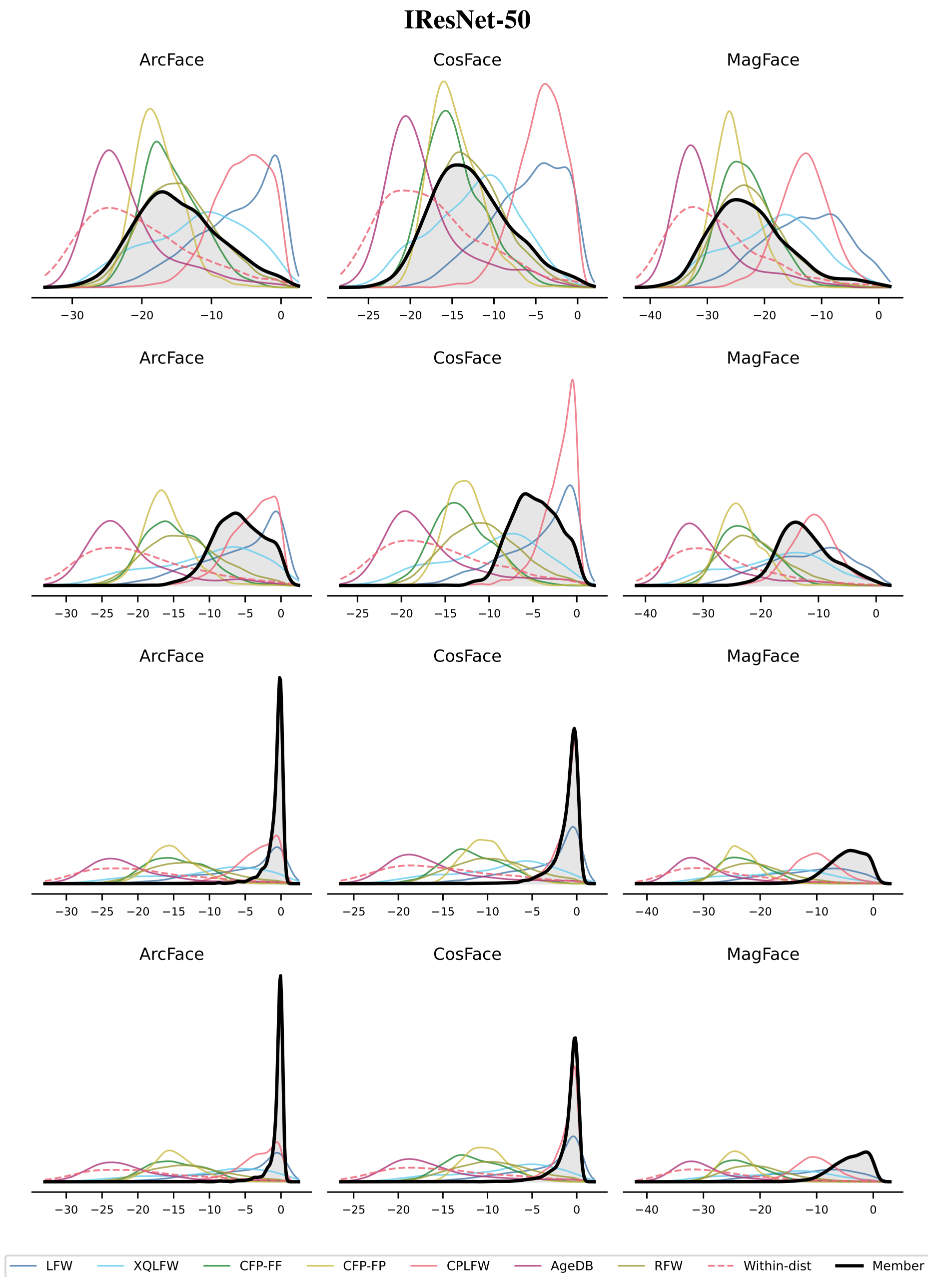


Figure 285. **Vary loss head — IResNet-50 / $n_{\text{ids}} = 1K$.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

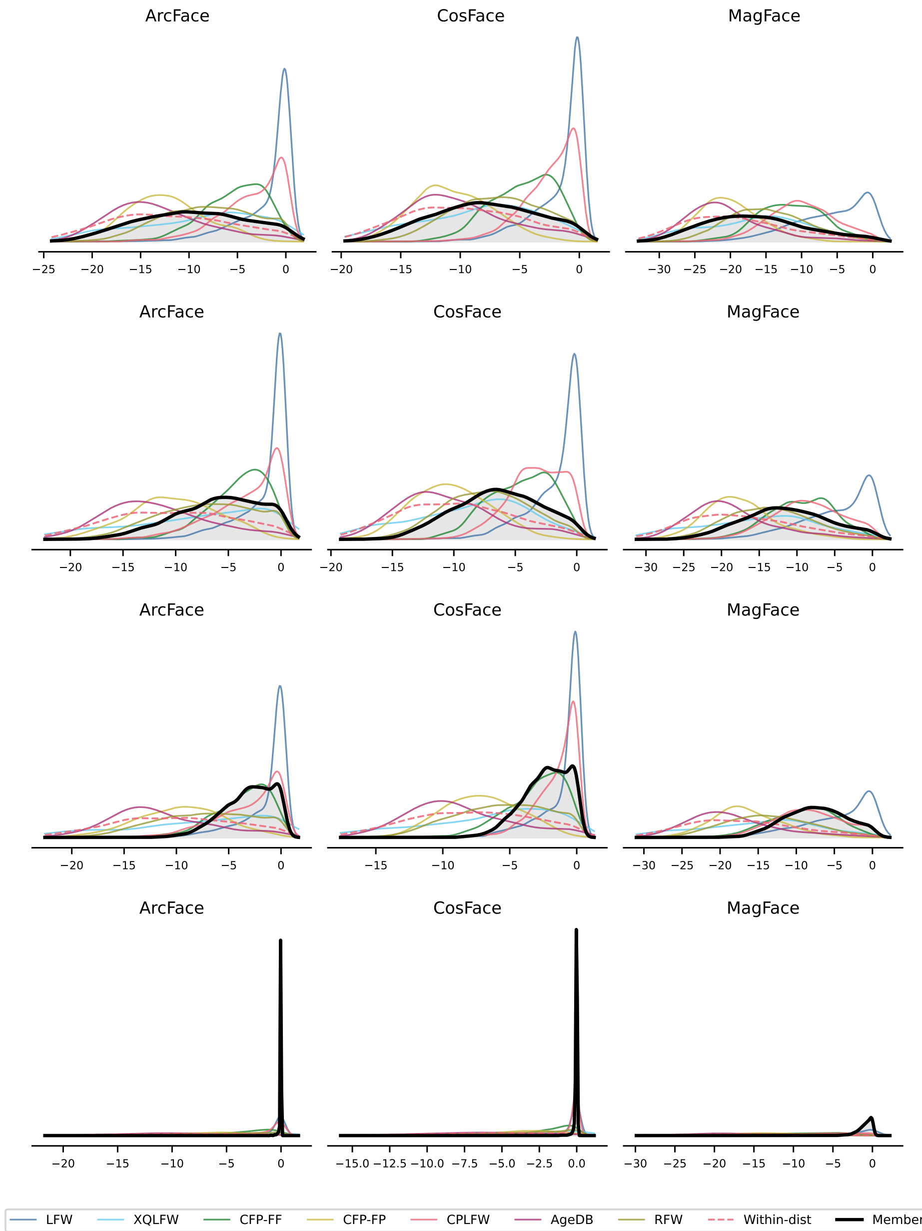


Figure 286. **Vary loss head — IResNet-50 /** $n_{\textbf{ids}} = 10K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

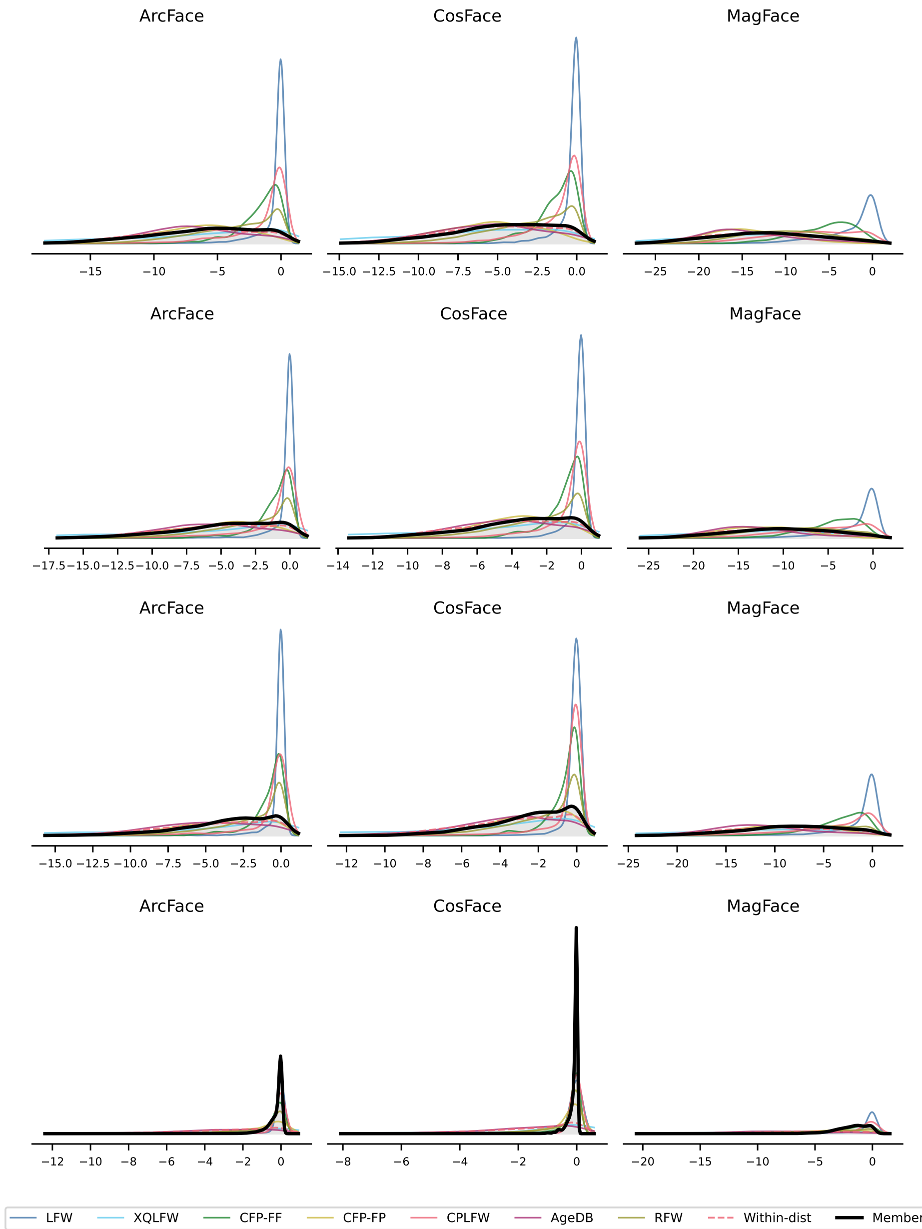


Figure 287. **Vary loss head — IResNet-50 /** $n_{\text{ids}} = 50K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

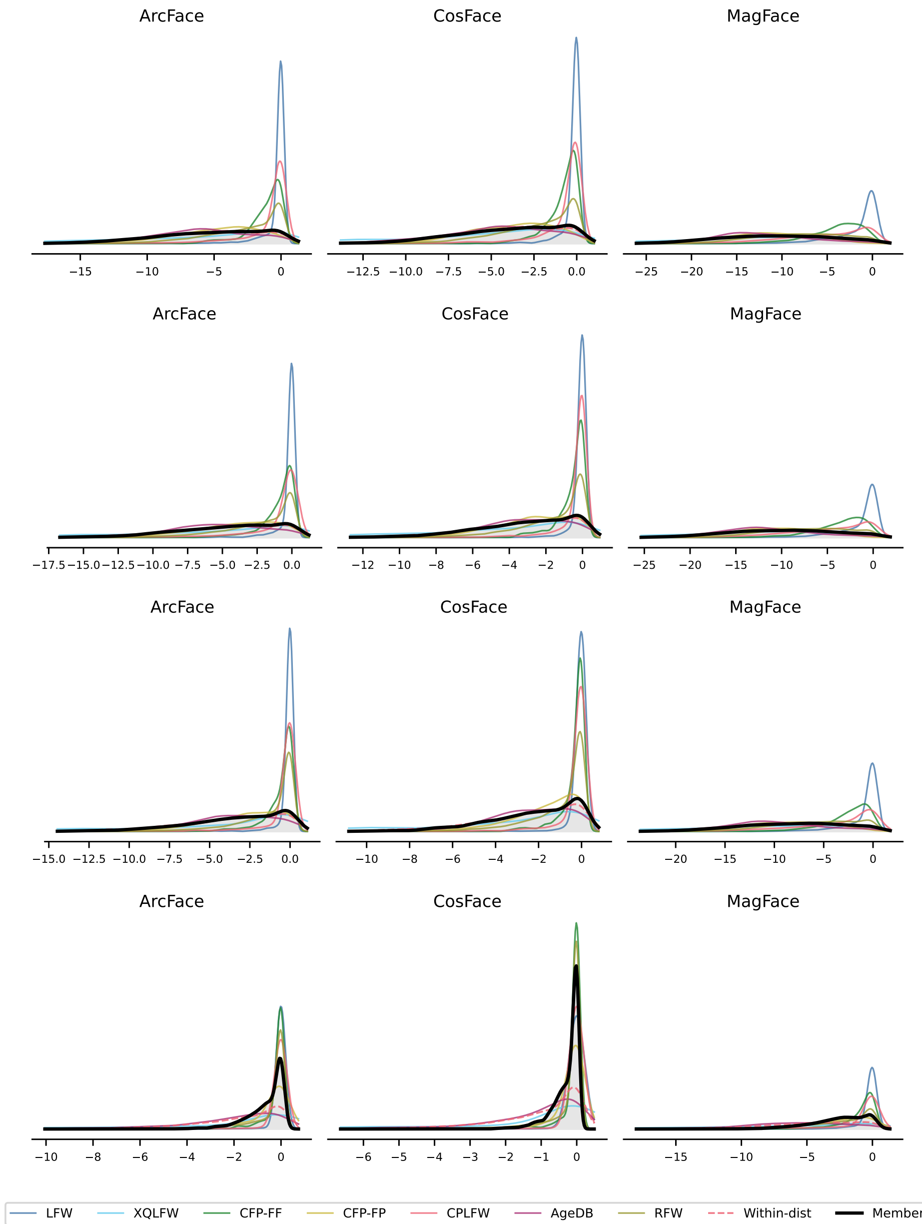


Figure 288. **Vary loss head — IResNet-50 /** $n_{\textbf{ids}} = 100K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

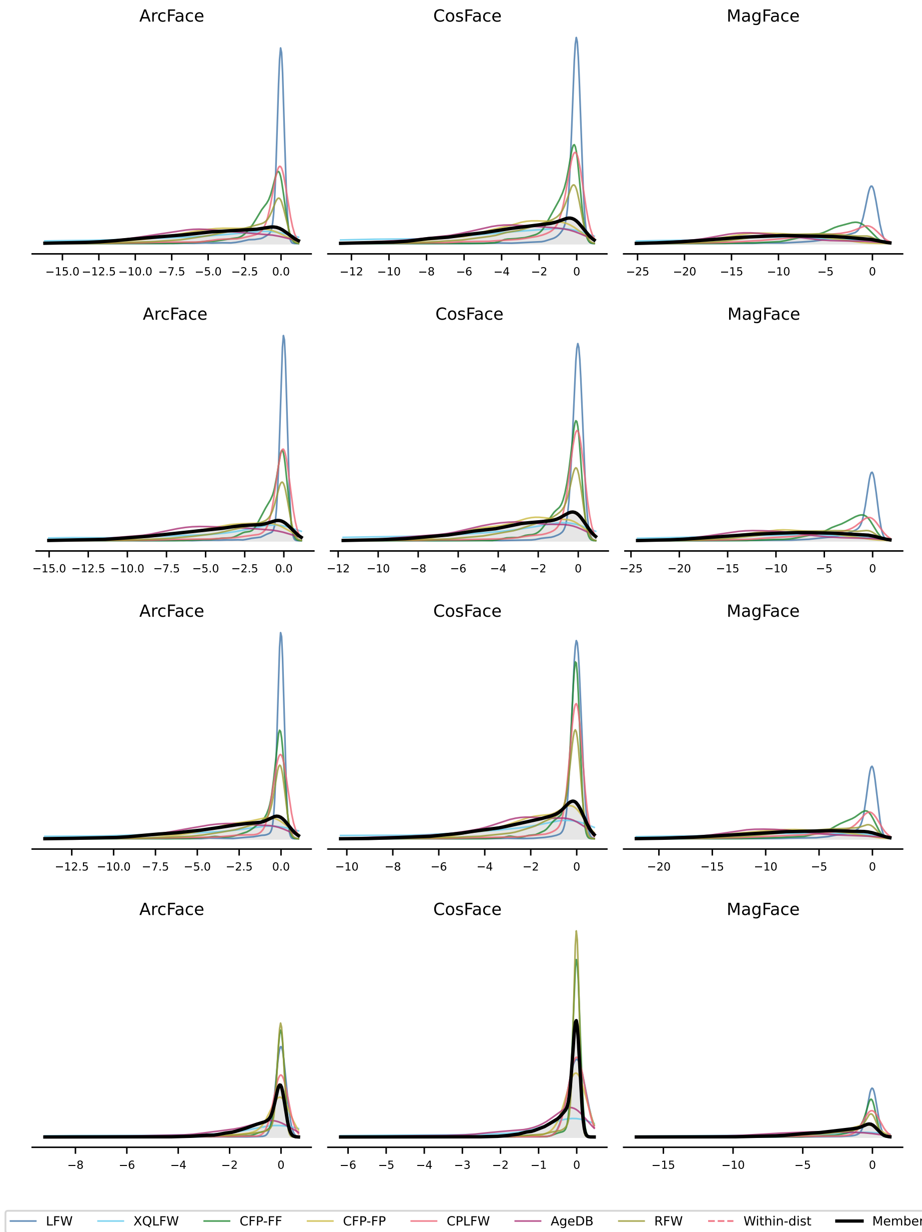


Figure 289. **Vary loss head — IResNet-50 /** $n_{\mathbf{ids}} = All$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

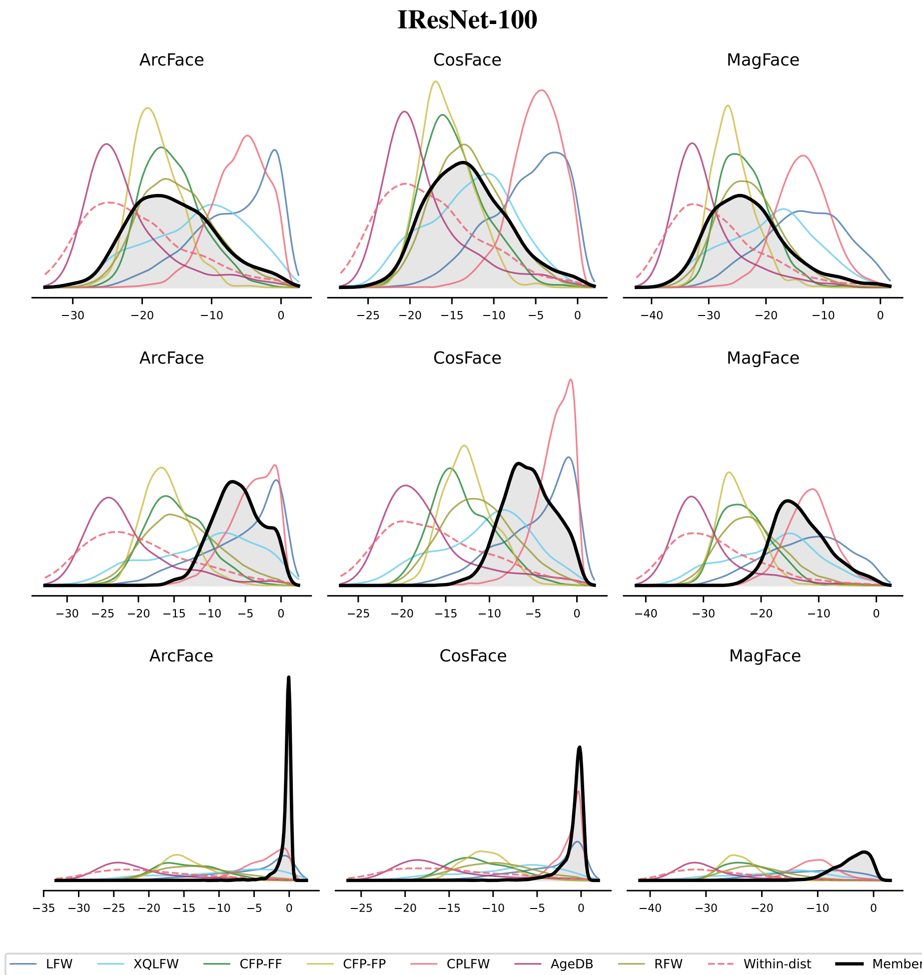


Figure 290. **Vary loss head — IResNet-100 / $n_{\text{ids}} = 1K$.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

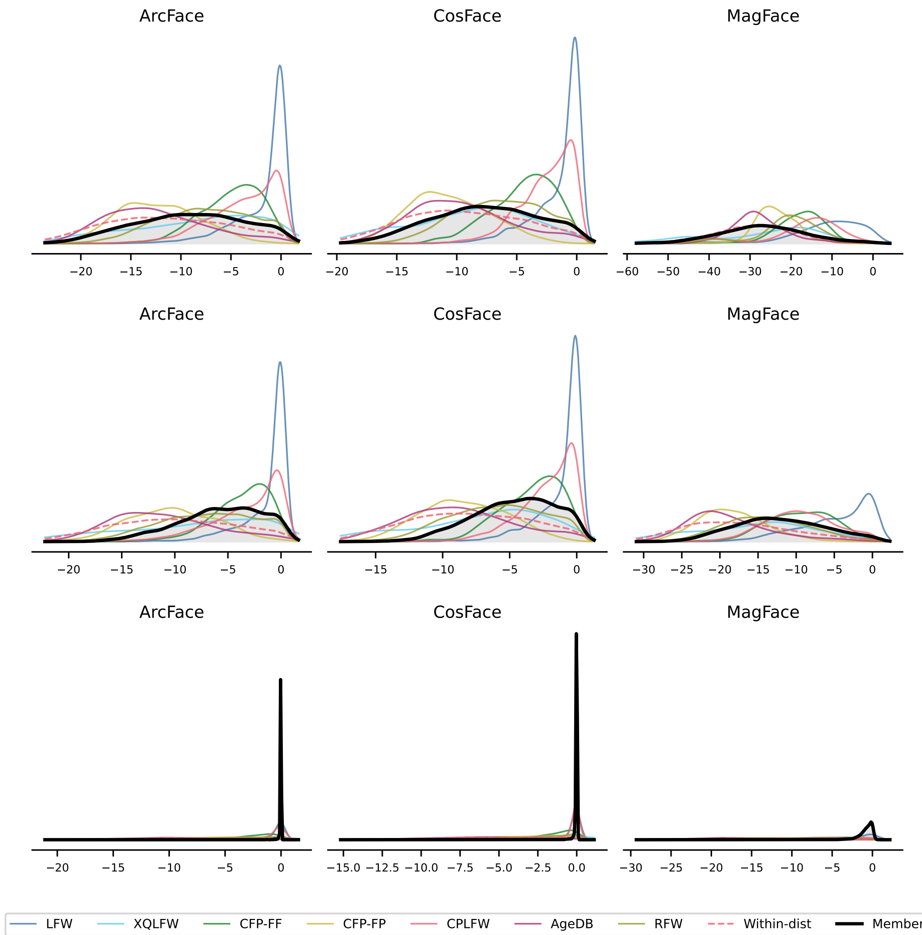


Figure 291. **Vary loss head — IResNet-100 /** $n_{\text{ids}} = 10K$**.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

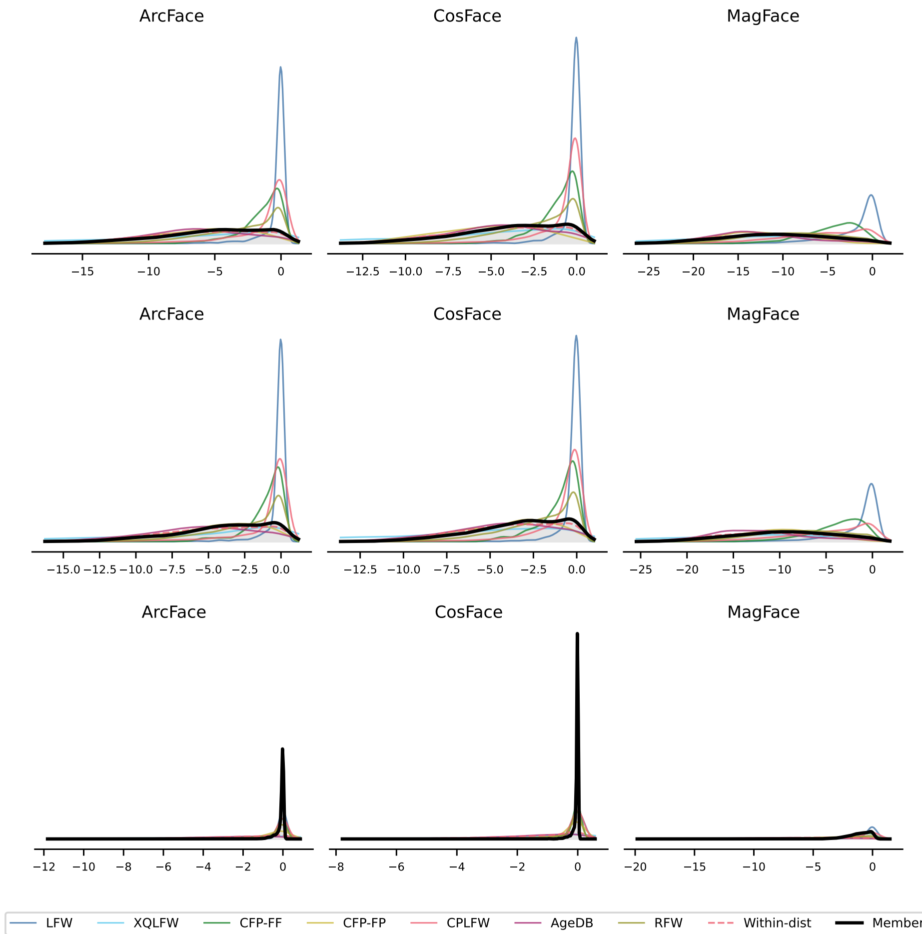


Figure 292. **Vary loss head — IResNet-100 /** $n_{\text{ids}} = 50K$**.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

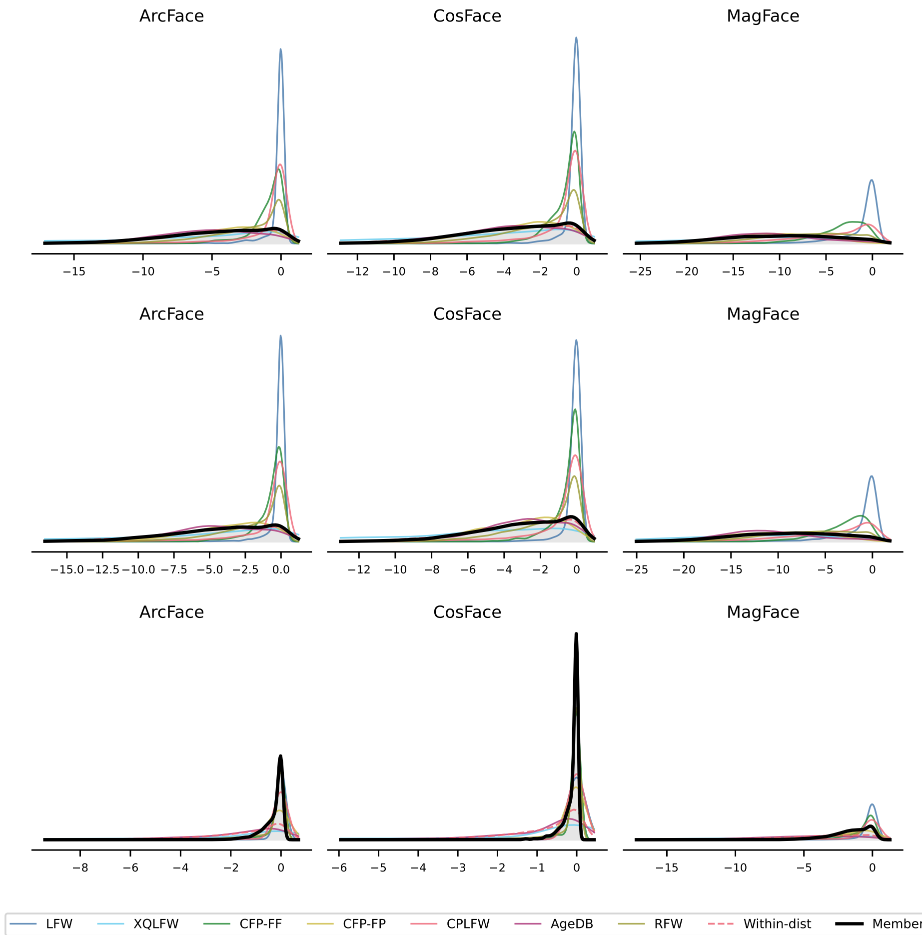


Figure 293. **Vary loss head — IResNet-100 /** $n_{\text{ids}} = 100K$**.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

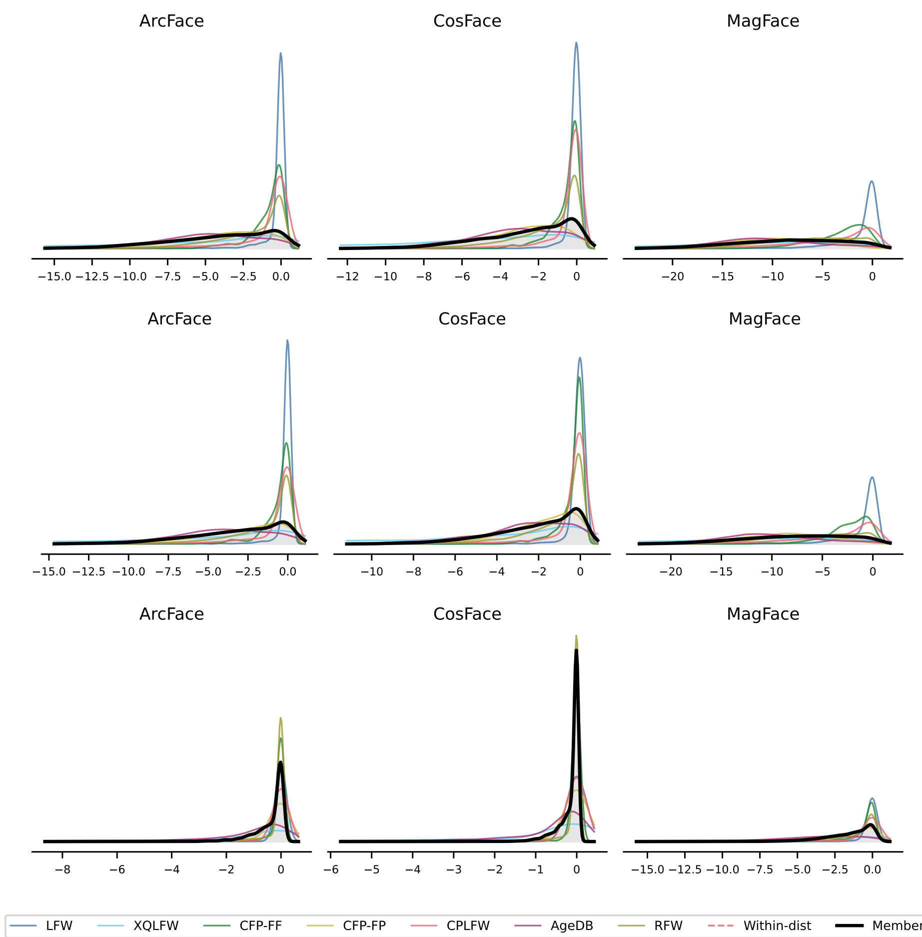


Figure 294. **Vary loss head — IResNet-100 / $n_{\text{ids}} = All$.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, ArcFace, CosFace, MagFace (left to right).

7.4. Vary backbone

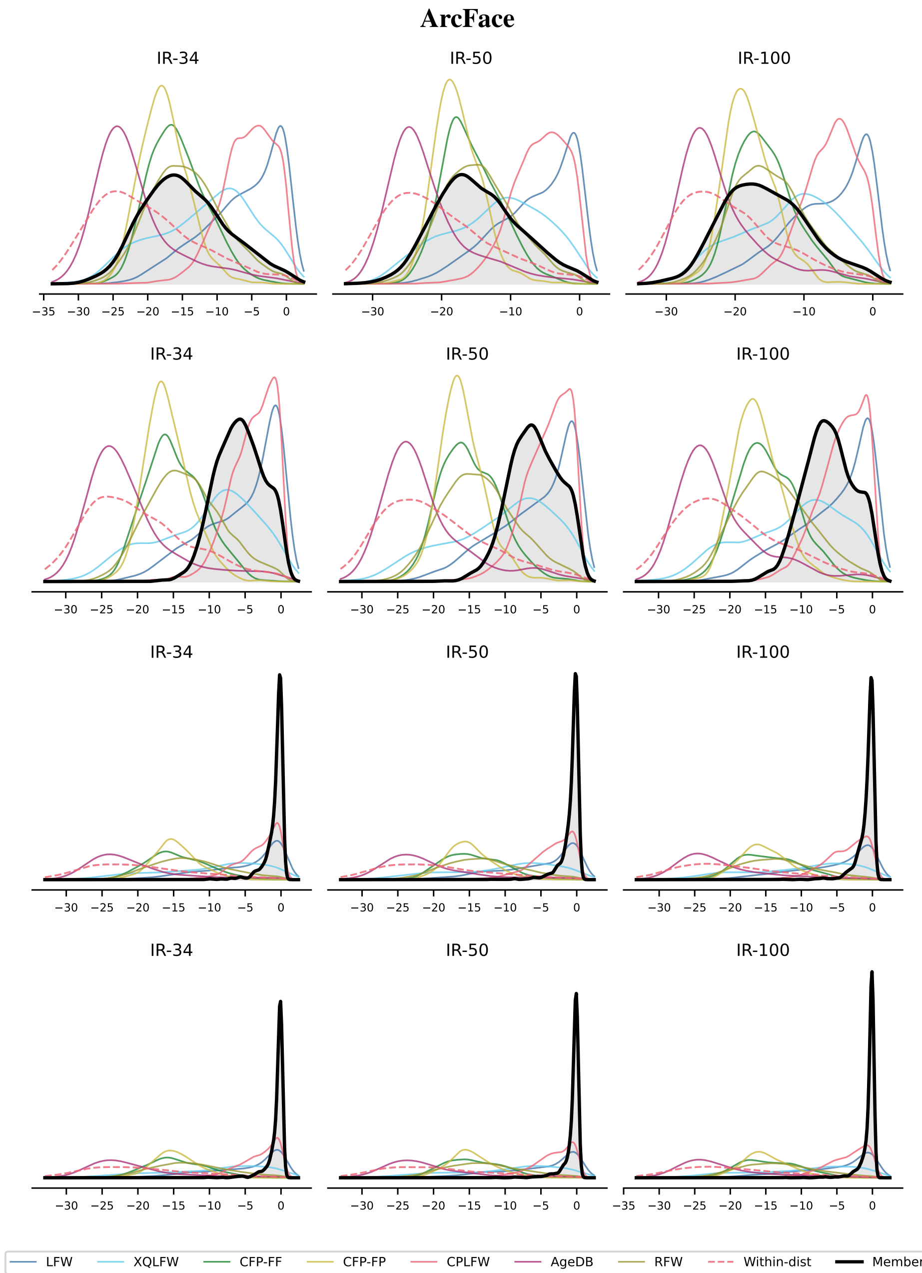


Figure 295. **Vary backbone — ArcFace /** $n_{\text{ids}} = 1K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

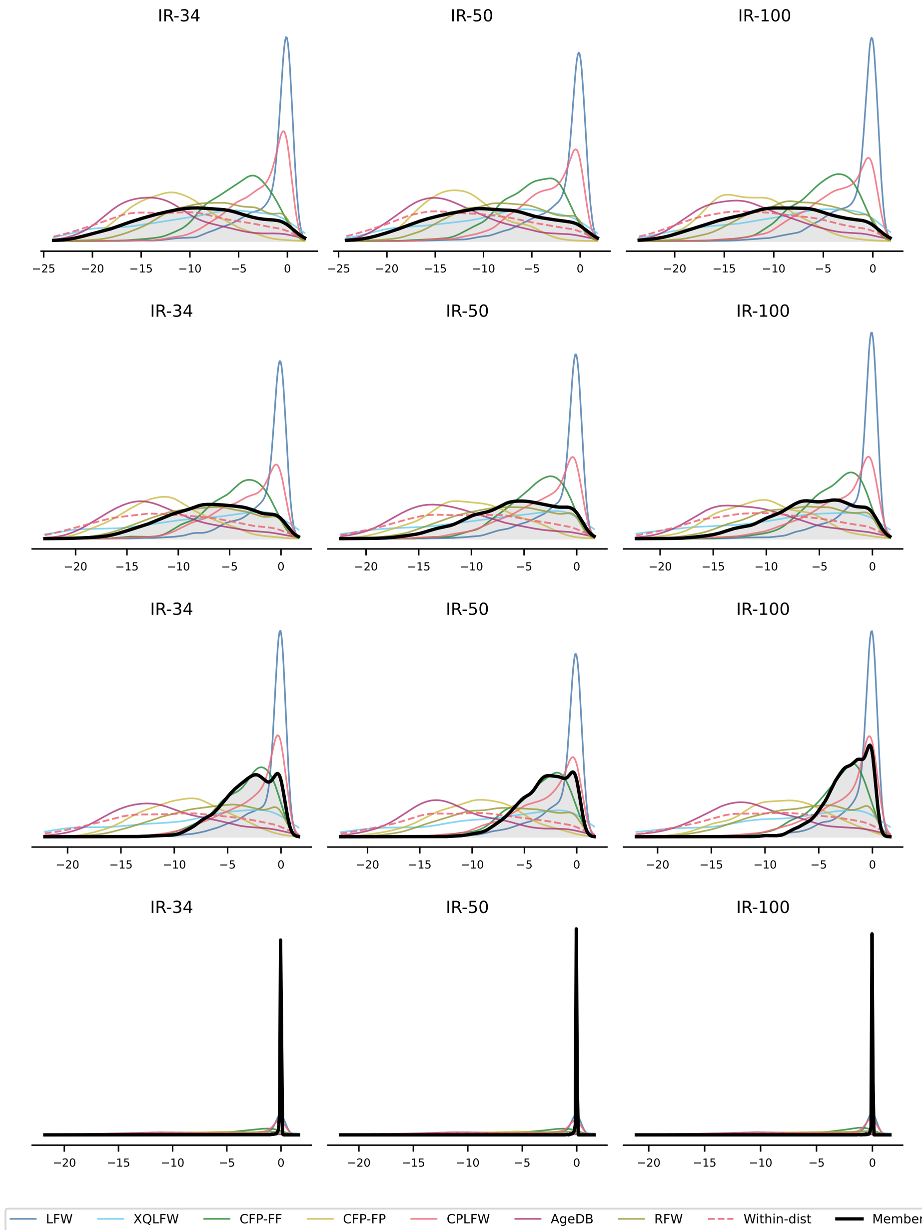


Figure 296. **Vary backbone — ArcFace /** $n_{\text{ids}} = 10K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

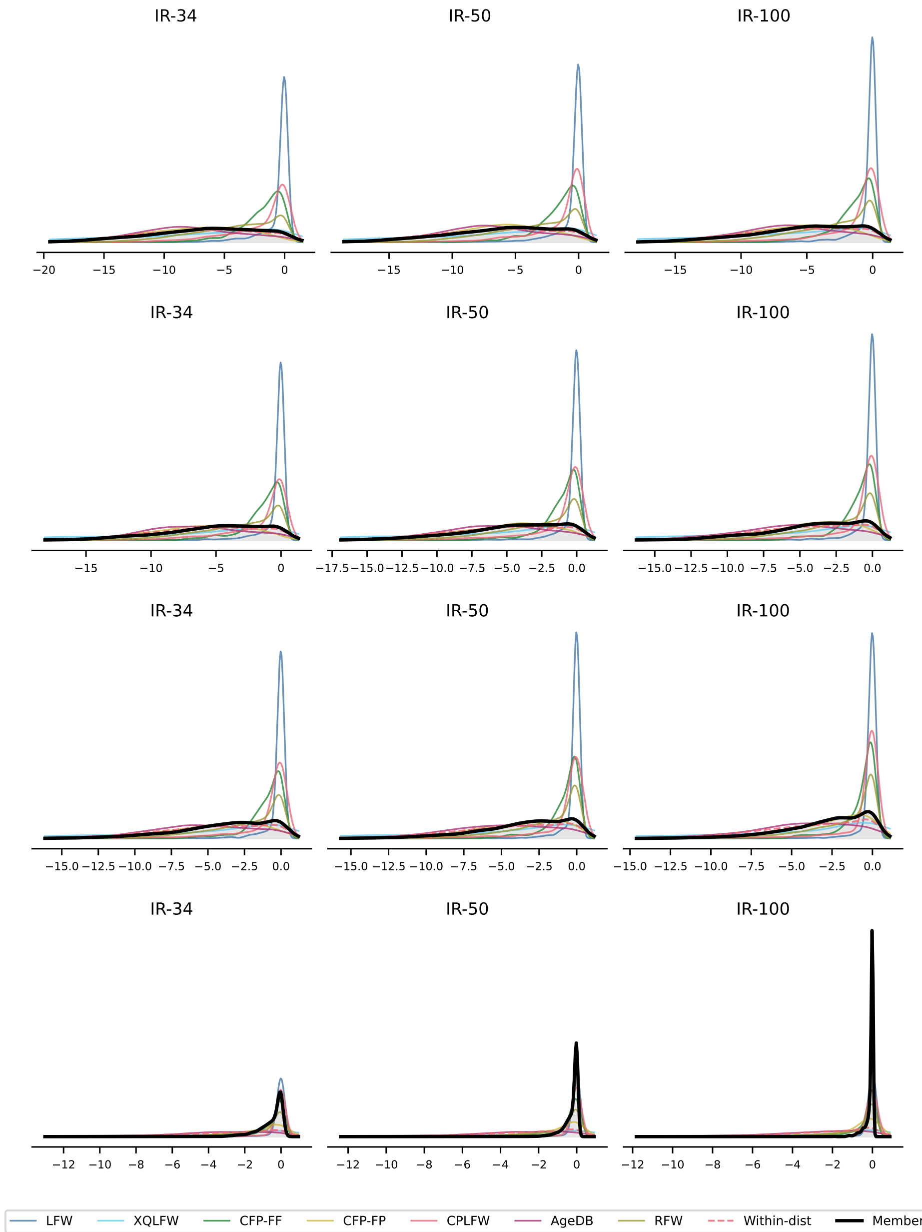


Figure 297. **Vary backbone — ArcFace /** $n_{\text{ids}} = 50K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

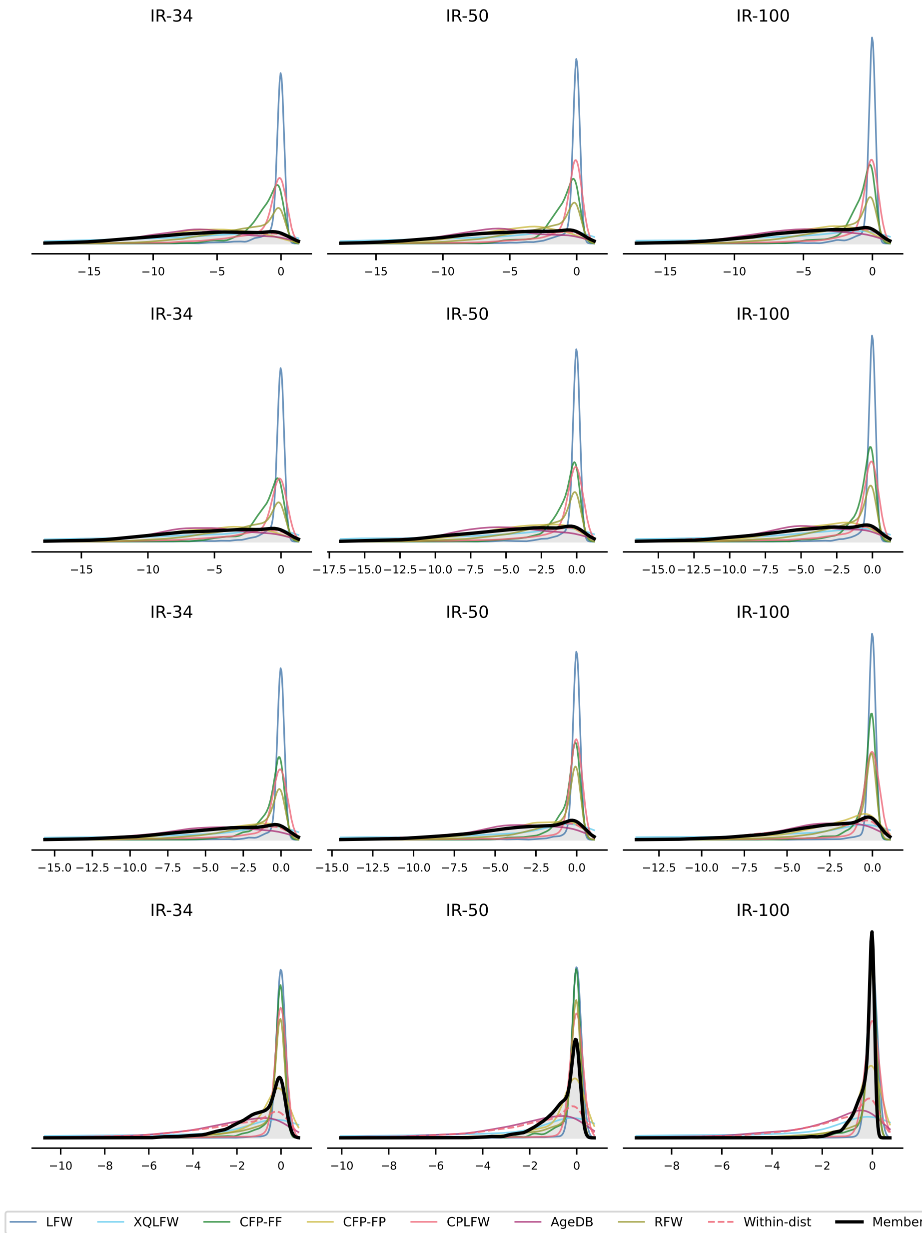


Figure 298. **Vary backbone — ArcFace / $n_{\text{ids}} = 100K$.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

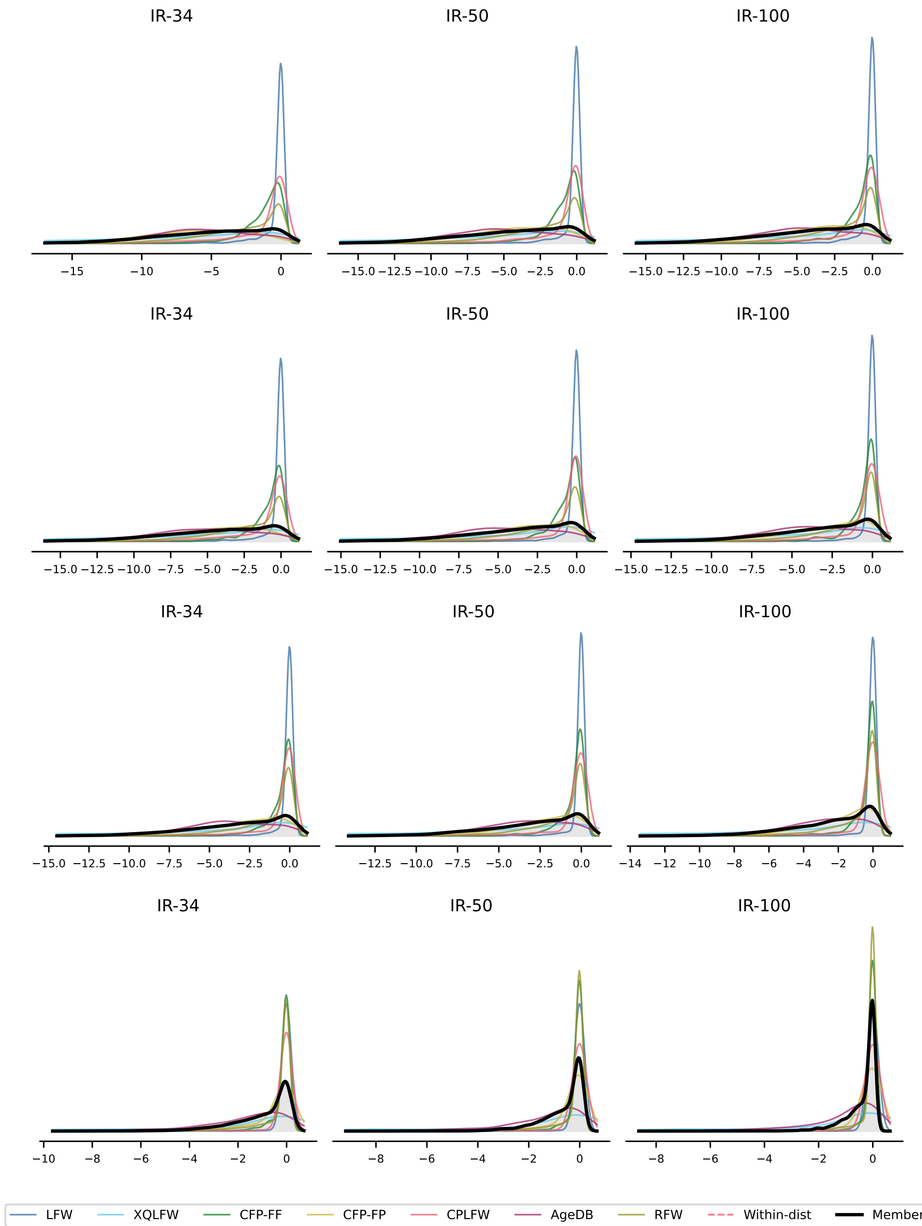


Figure 299. **Vary backbone — ArcFace /** $n_{\mathbf{ids}} = All$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

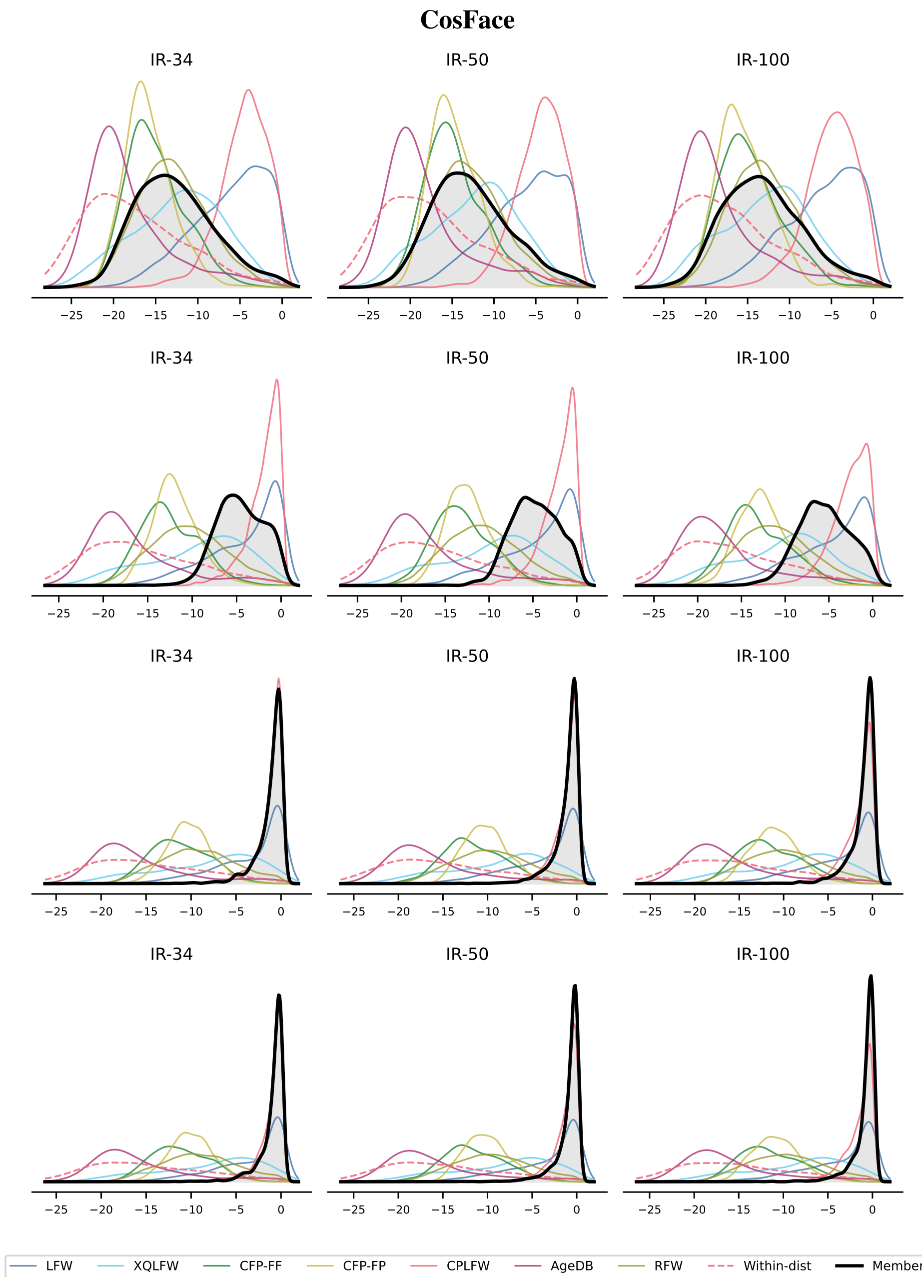


Figure 300. **Vary backbone — CosFace / $n_{\text{ids}} = 1K$.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

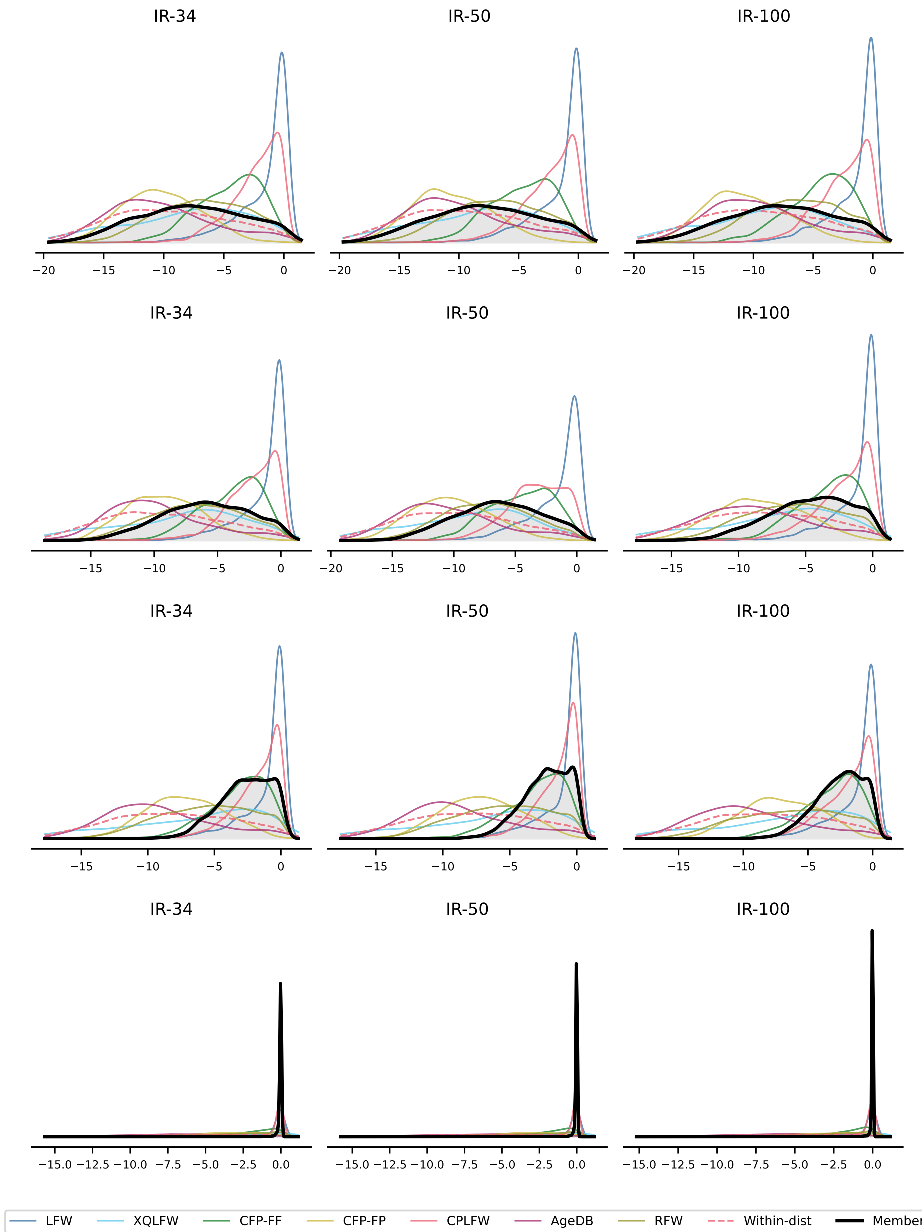


Figure 301. **Vary backbone — CosFace /** $n_{\text{ids}} = 10K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

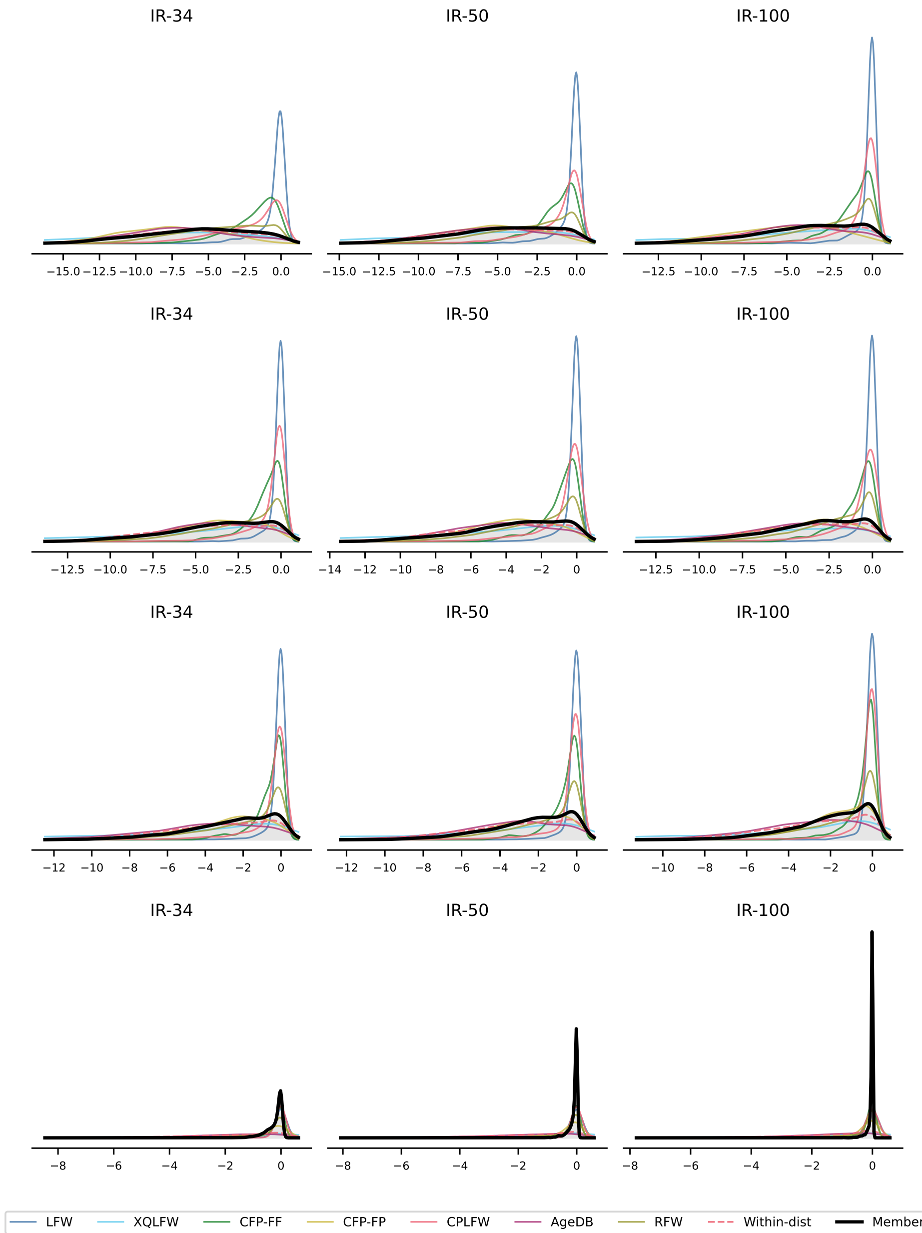


Figure 302. **Vary backbone — CosFace /** $n_{\text{ids}} = 50K$**.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

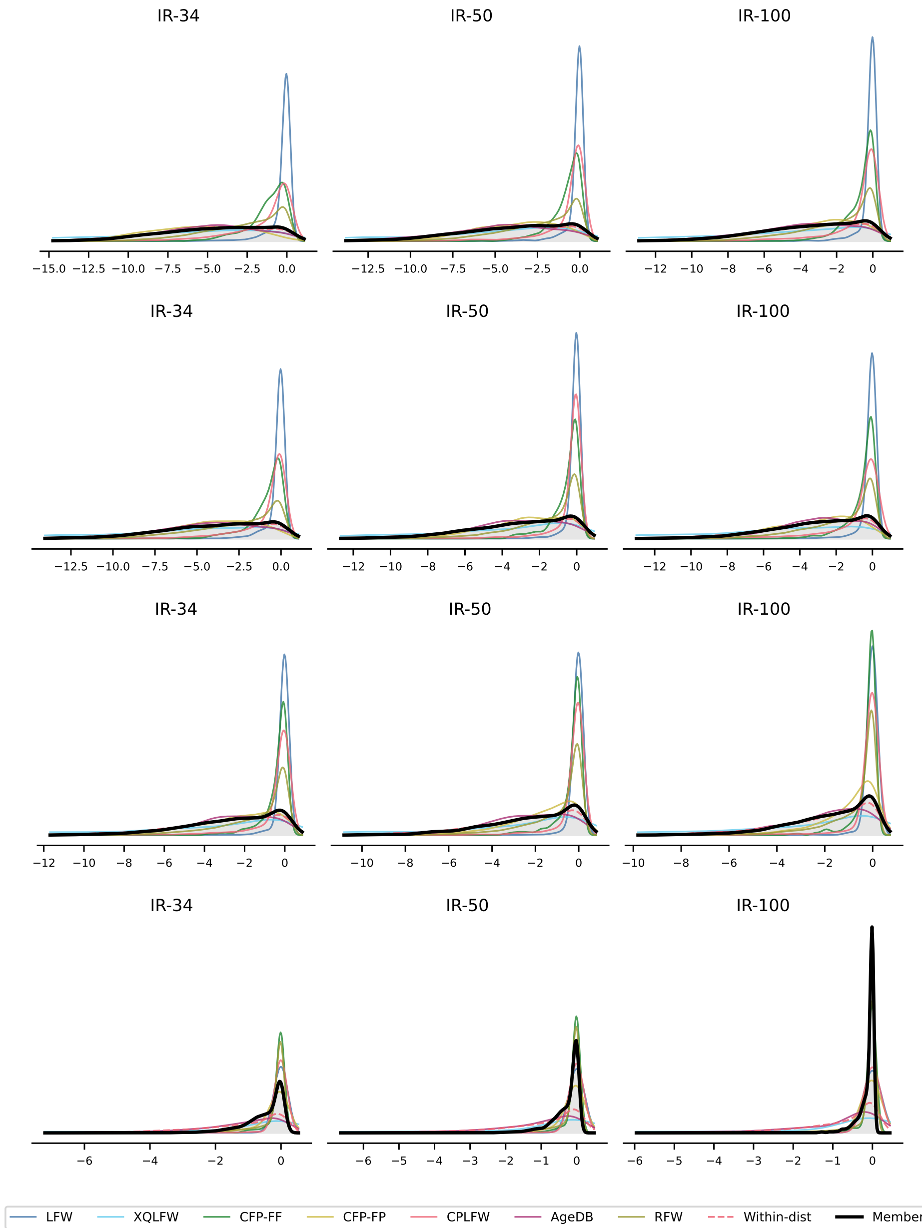


Figure 303. **Vary backbone — CosFace / $n_{\text{ids}} = 100K$.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

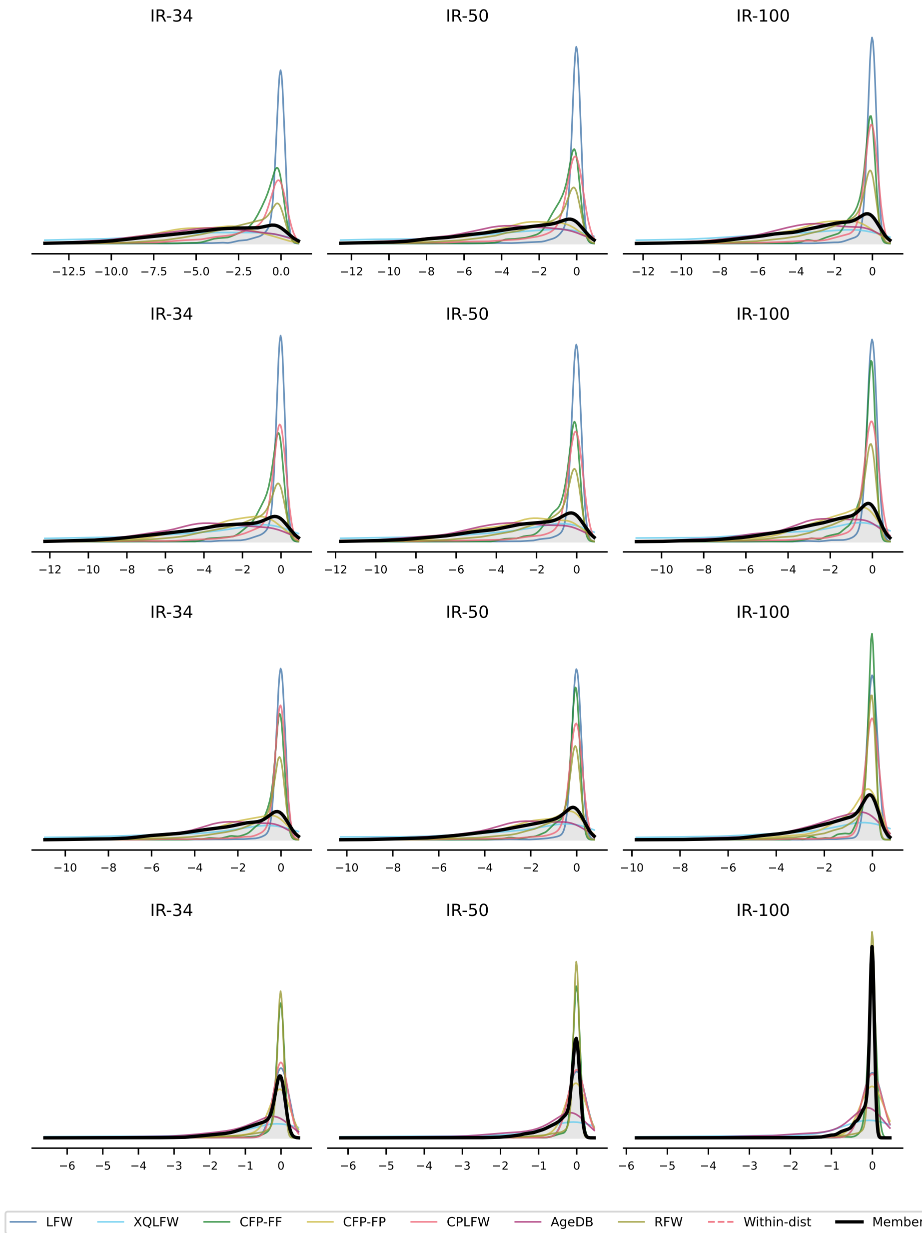


Figure 304. **Vary backbone — CosFace / $n_{\text{ids}} = All$.** Rows: epochs 5, 10, 20, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

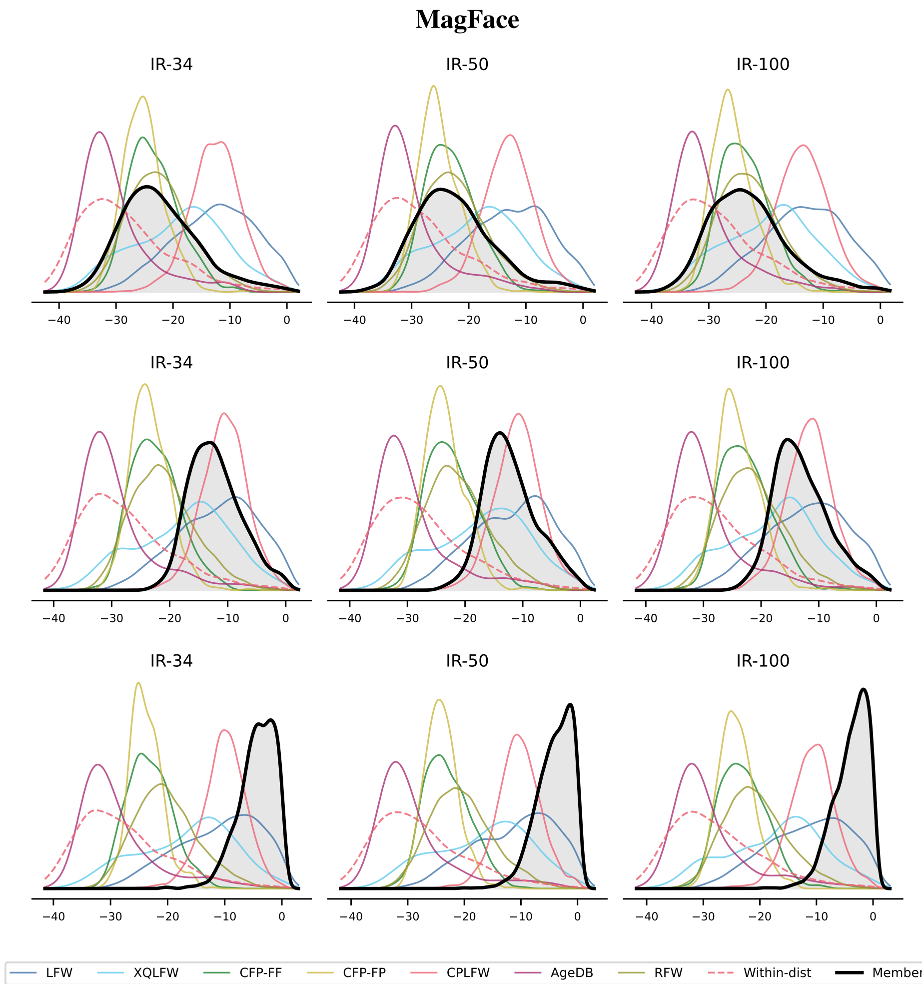


Figure 305. **Vary backbone — MagFace /** $n_{\text{ids}} = 1K$**.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

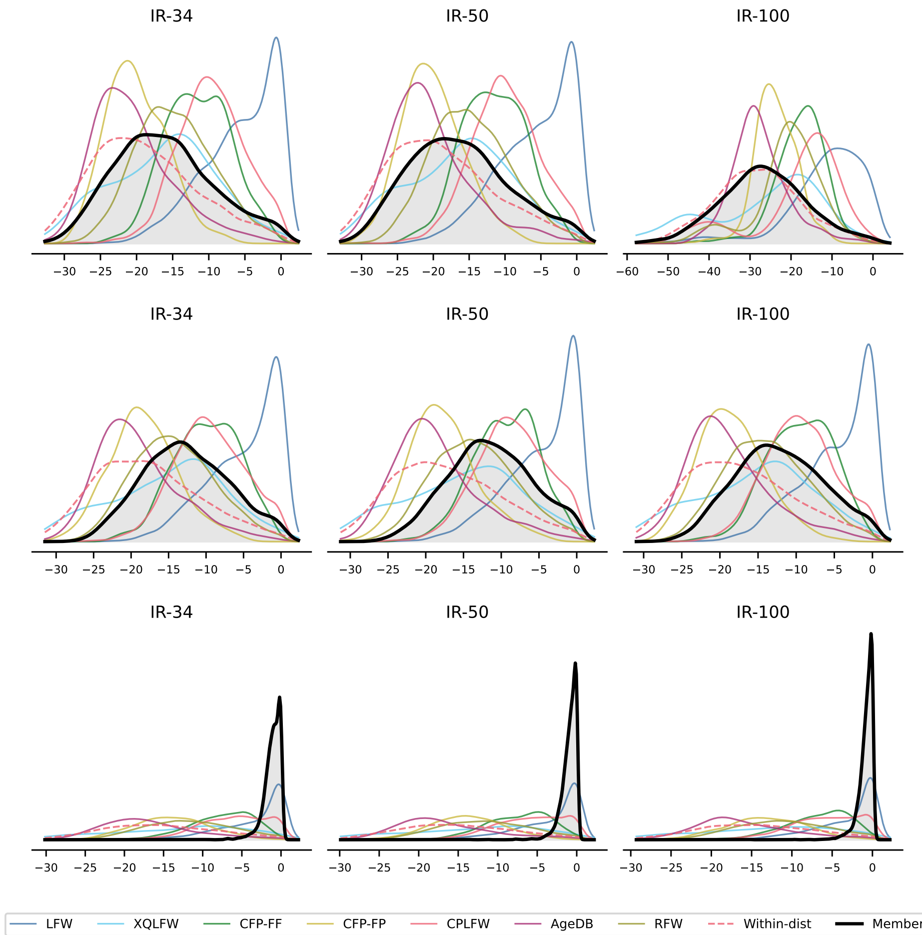


Figure 306. **Vary backbone — MagFace /** $n_{\text{ids}} = 10K$**.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

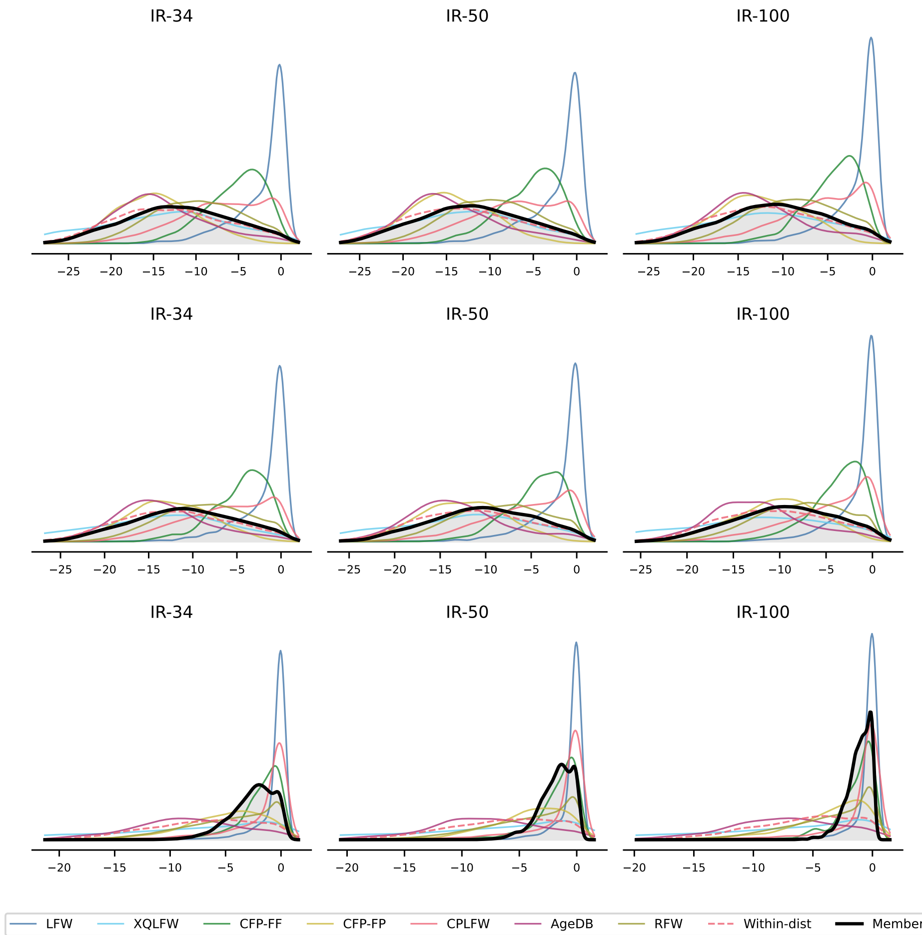


Figure 307. **Vary backbone — MagFace /** $n_{\textbf{ids}} = 50K$**.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

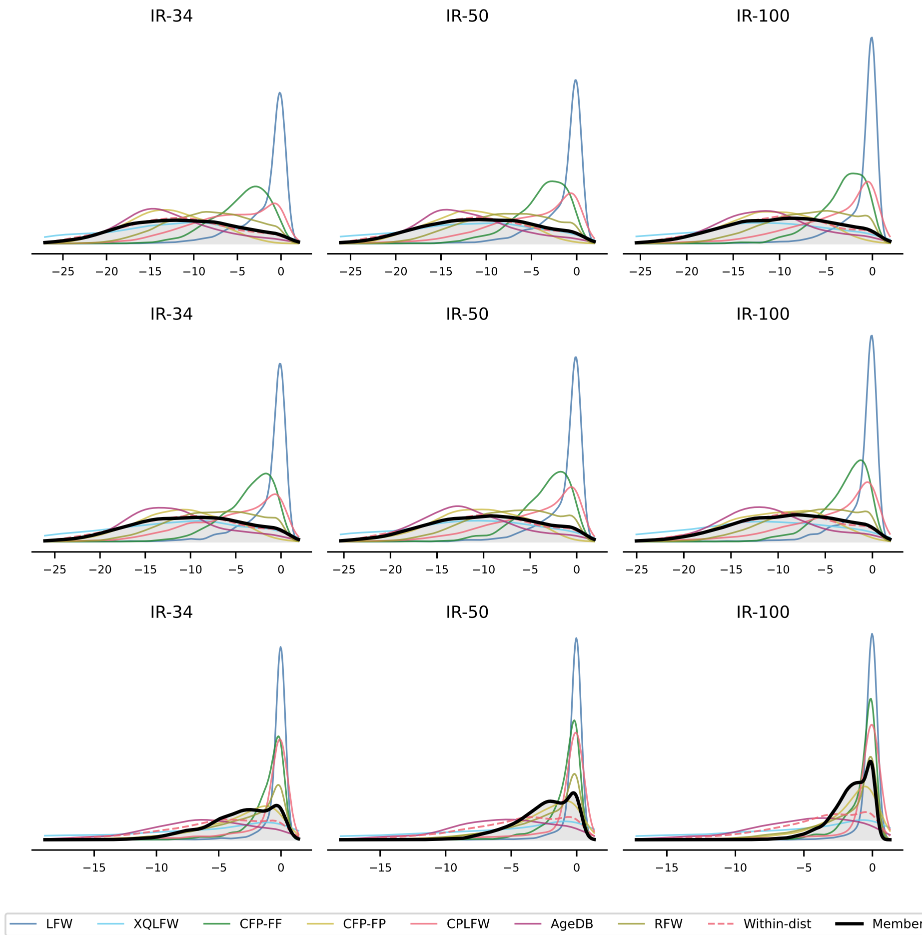


Figure 308. **Vary backbone — MagFace /** $n_{\mathbf{ids}} = 100K$**.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).

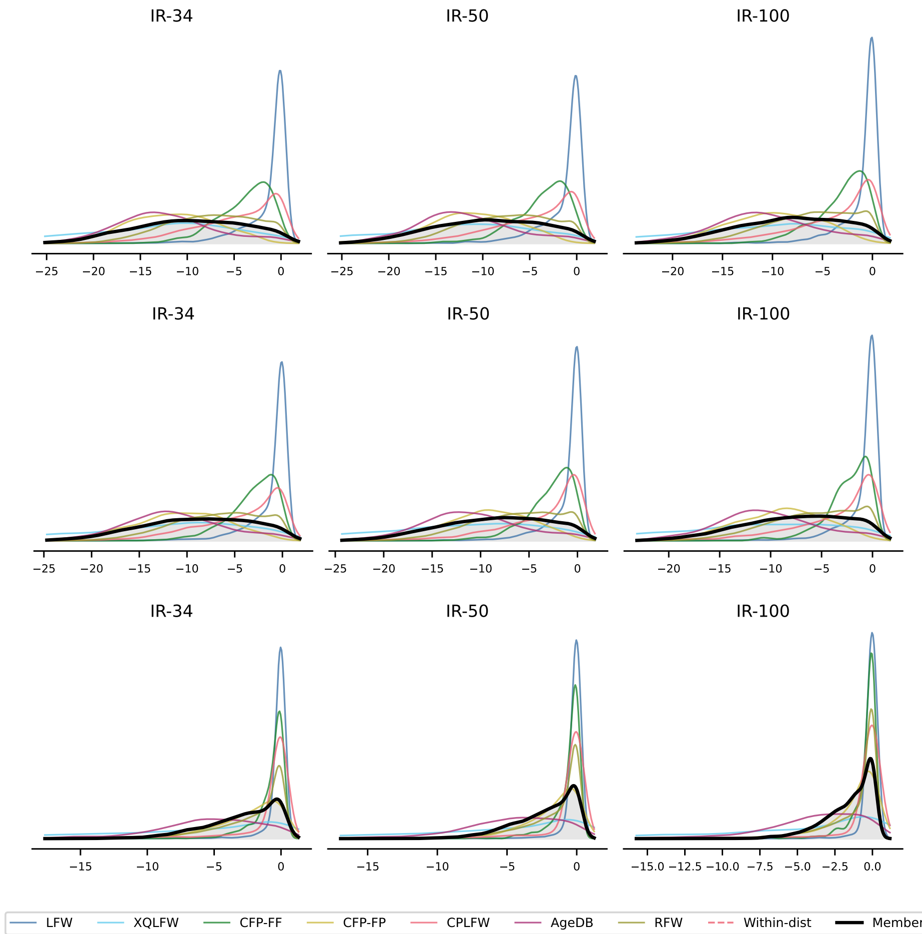


Figure 309. **Vary backbone — MagFace /** $n_{\text{ids}} = All$**.** Rows: epochs 5, 10, and 40 (top to bottom). Within each row, IResNet-34, IResNet-50, IResNet-100 (left to right).