A. Proof of Equation 6

Here, we give details of the derivation of the $\mathbf{y}$-updates in Eq. (6) from the upper bound (majorizing function) in (9). Given a solution $\mathbf{y}^{(n)}$ at iteration n, the goal is to find the next iterate $\mathbf{y}^{(n+1)}$ that minimizes the following tight upper bound, s.t. simplex constraint $\mathbf{y} \in \Delta^{N-1}$:

$$\mathcal{B}(\mathbf{y}, \mathbf{y}^{(n)}) = -\mathbf{y}^t\mathbf{k} - \frac{\lambda}{2}\mathbf{y}^t W^t \mathbf{y}^{(n)} - \lambda_{\mathbf{y}} H(\mathbf{y}) \tag{9}$$

where $\mathbf{k} = (K(\mathbf{f}_p - \mathbf{m}))_{1 \leq p \leq N}$.

The objective function of (9) is strictly convex. Taking into account the simplex constraint on $\mathbf{y}$, the associated Lagrangian reads:

$$\mathcal{L}(\mathbf{y}, \mathbf{y}^{(n)}) = -\mathbf{y}^t\mathbf{k} - \frac{\lambda}{2}\mathbf{y}^t W^t \mathbf{y}^{(n)} - \lambda_{\mathbf{y}} H(\mathbf{y}) + \gamma(\mathbf{y}^t \mathbf{1}_N - 1) \tag{10}$$

where γ is the Lagrange multiplier for simplex constraint $\mathbf{y} \in \Delta^{N-1}$ and $\mathbf{1}_N$ is the vector of ones. Note that we do not impose explicitly the constraints on the non-negativity of the components of $\mathbf{y}$ because these are implicitly enforced with the entropic barrier term in (9), i.e., $-\lambda_{\mathbf{y}} H(\mathbf{y})$. Now, computing the gradient of $\mathcal{L}(\mathbf{y}, \mathbf{y}^{(n)})$ w.r.t $\mathbf{y}$ yields:

$$\nabla_{\mathbf{y}}\mathcal{L}(\mathbf{y}, \mathbf{y}^{(n)}) = -\mathbf{k} - \lambda W^t \mathbf{y}^{(n)} + (\gamma + \lambda_{\mathbf{y}})\mathbf{1}_N + \lambda_{\mathbf{y}} \log(\mathbf{y}) \tag{11}$$

By setting the gradients of (11) to 0, we get the optimal solution:

$$y_p^{(n+1)} = \exp\left((K(\mathbf{f}_p - \mathbf{m}) + \lambda \sum_{q=1}^{N} w_{p,q} y_q^{(n)})/\lambda_{\mathbf{y}} \right) . \\ \exp\left(-(\gamma + \lambda_{\mathbf{y}})\right) \tag{12}$$

Using this expression in the simplex constraint $\sum_p y_p^{(n+1)} = 1$ enables to recover the following expression of $\exp\left(\gamma + \lambda_{\mathbf{y}}\right)$:

$$\sum_{j=1}^{N} \exp\left((K(\mathbf{f}_j - \mathbf{m}) + \lambda \sum_{q=1}^{N} w_{j,q} y_q^{(n)})/\lambda_{\mathbf{y}} \right)$$

Plugging this expression back in (12), we get the final updates:

$$y_p^{(n+1)} = \frac{\exp\left((K(\mathbf{f}_p - \mathbf{m}) + \lambda \sum_{q=1}^{N} w_{p,q} y_q^{(n)})/\lambda_{\mathbf{y}} \right)}{\sum_{j=1}^{N} \exp\left((K(\mathbf{f}_j - \mathbf{m}) + \lambda \sum_{q=1}^{N} w_{j,q} y_q^{(n)})/\lambda_{\mathbf{y}} \right)} \tag{13}$$

B. Cauchy and convergent sequence proof

Let us consider iteration l, with the associated current *inlierness* scores $\mathbf{y}$. Let us prove that $\{\mathbf{m}^l\}_{l \in \mathbb{N}}$ is a Cauchy sequence. Recall the recursive relation:

$$\mathbf{m}^{l+1} = \frac{\sum_{p=1}^{N} y_p K(\mathbf{f}_p - \mathbf{m}^l)\mathbf{f}_p}{\sum_{p=1}^{N} y_p K(\mathbf{f}_p - \mathbf{m}^l)} \tag{14}$$

with $K(\mathbf{f}_p - \mathbf{m}^l) = \exp(-\|\frac{\mathbf{f}_p - \mathbf{m}^l}{h}\|^2)$, for some $h > 0$. We define:

$$k(x) = \exp(-x) \tag{15}$$

$$u^l = \sum_{p=1}^{N} y_p K(\mathbf{f}_p - \mathbf{m}^l) \tag{16}$$

$$v^l = \sum_{p=1}^{N} y_p K(\mathbf{f}_p - \mathbf{m}^l)\mathbf{f}_p \tag{17}$$

Step 1: First, let us prove that $\{u^l\}_{l \in \mathbb{N}}$ is a Cauchy sequence. Recall that in a metric space, a convergent sequence is necessarily a Cauchy sequence. Therefore, we only need to show that u^l is convergent (i.e bounded and strictly monotonic).

Notice that for $x > 0$, $0 \leq k(x) \leq 1$. Therefore:

$$u^l = \sum_{p=1}^{N} y_p k(\|\frac{\mathbf{f}_p - \mathbf{m}^l}{h}\|^2) \tag{18}$$

$$\leq \sum_{p=1}^{N} y_p \leq 1 \tag{19}$$

Therefore, u^l is bounded between 0 and 1. Now, let us study the consecutive differences $\Delta^l = u^{l+1} - u^l$:

$$\Delta^l = \sum_{p=1}^{N} y_p \left[k(\frac{\|\mathbf{f}_p - \mathbf{m}^{l+1}\|^2}{h^2}) - k(\frac{\|\mathbf{f}_p - \mathbf{m}^l\|^2}{h^2}) \right] \tag{20}$$

Because k is convex, one can say that $\forall a, b \in \mathbb{R}$:

$$k(a) - k(b) \geq k'(b)(a - b) \tag{21}$$

And because $k'(b) = -k(b)$ in our case, one ends up with:

$$k(a) - k(b) \geq k(b)(b - a) \tag{22}$$

Applied with $a = \frac{\|\mathbf{f}_p - \mathbf{m}^{l+1}\|^2}{h^2}$ and $b = \frac{\|\mathbf{f}_p - \mathbf{m}^l\|^2}{h^2}$, one can obtain:

$$\Delta^l \geq \sum_{p=1}^{N} y_p K(\mathbf{f}_p - \mathbf{m}^l) \left[\frac{\|\mathbf{f}_p - \mathbf{m}^l\|^2}{h^2} - \frac{\|\mathbf{f}_p - \mathbf{m}^{l+1}\|^2}{h^2} \right]$$

$$= \frac{1}{h^2} \sum_{p=1}^{N} y_p K(\mathbf{f}_p - \mathbf{m}^l) \left[\|\mathbf{m}^l\|^2 - \|\mathbf{m}^{l+1}\|^2 \right. \\ \left. -2 < \mathbf{m}^l, \mathbf{f}_p > +2 < \mathbf{m}^{l+1}, \mathbf{f}_p > \right]$$

Now is time to recall recursive relation $\mathbf{m}^{l+1} = \frac{v^l}{u^l}$. By simply expanding, one can end up with:

$$\begin{aligned}\Delta^l \geq &\frac{1}{h^2}\left[\|\mathbf{m}^l\|^2 u^l - \frac{\|v^l\|^2}{u^l} - 2 < \mathbf{m}^l, v^l > +2\frac{\|v^l\|^2}{u^l}\right] \\ = &\frac{1}{h^2} u^l \left[\|\mathbf{m}^l\|^2 - 2 < \mathbf{m}^l, \mathbf{m}^{l+1} > + \|\mathbf{m}^{l+1}\|^2\right] \\ = &\frac{1}{h^2} u^l \|\mathbf{m}^l - \mathbf{m}^{l+1}\|^2 \qquad (23)\end{aligned}$$

Therefore, $\Delta^l > 0$, which shows that $\{u^l\}_{l\in\mathbb{N}}$ is strictly increasing. This concludes the proof that u^l is a convergent sequence, and therefore a Cauchy one.

Step 2: Now, on top of concluding the proof that $\{u^l\}_{l\in\mathbb{N}}$ is a Cauchy sequence, Eq. (23) also offers an interesting relation between $\{\Delta^l\}_{l\in\mathbb{N}}$ and the sequence of interest $\{\mathbf{m}^l\}_{l\in\mathbb{N}}$, which we can use. Indeed, for any $l_0, m \in \mathbb{N}$, we can sum Eq. (23):

$$\sum_{l=l_0}^{l_0+m} \Delta^l \geq \frac{1}{h^2} \sum_{l=l_0}^{l_0+m} u^l \|\mathbf{m}^l - \mathbf{m}^{l+1}\|^2 \qquad (24)$$

$$\geq \frac{u^{l_0}}{h^2} \sum_{l=l_0}^{l_0+m} \|\mathbf{m}^l - \mathbf{m}^{l+1}\|^2 \qquad (25)$$

$$\geq \frac{u^{l_0}}{h^2} \|\mathbf{m}^{l_0+m} - \mathbf{m}^{l_0}\|^2 \qquad (26)$$

Where Eq. (25) follows because $\{u^l\}_{l\in\mathbb{N}}$ is strictly increasing, and Eq. (26) follows from the triangle inequality. Now, the left-hand side of Eq. (24) can be reduced to $\sum_{l=l_0}^{l_0+m} \Delta^l = u^{l_0+m+1} - u^{l_0}$. But because we proved in Step 1 that $\{u^l\}_{l\in\mathbb{N}}$ was a Cauchy sequence, this difference is bounded by a constant. This concludes the proof that $\{\mathbf{m}^l\}_{l\in\mathbb{N}}$ is itself a Cauchy sequence in the Euclidean space.

Step 3: We just proved that $\{\mathbf{m}^l\}_{l\in\mathbb{N}}$ was a Cauchy sequence. Therefore $\{\mathbf{m}^l\}_{l\in\mathbb{N}}$ can only converge to a single value $\mathbf{m}^*$. We now use the continuity of function g to conclude that $\mathbf{m}^*$ has to be a solution of the initial equation (7):

$$\mathbf{m}^* = \lim_{l\to\infty} \mathbf{m}^{l+1} = \lim_{l\to\infty} g(\mathbf{m}^l) \qquad (27)$$

$$= g(\lim_{l\to\infty} \mathbf{m}^l) = g(\mathbf{m}^*) \qquad (28)$$

C. Further details on MTA

We summarize the traditional mode seeking MeanShift procedure, upon which our approach is based, in Algorithm 1. Moreover, our robust multi-modal MeanShift for test-time augmentation, named MTA, is presented in Algorithm 2 in a non-vectorized manner to highlight each operation. The handcrafted prompts [46] for ensembling are listed in Table 12. We use N=64 augmented views (63 from random cropping (RandomCrop) and the original image) in all our experiments except in Table 3 to be consistent with DiffTPT which uses 128 augmented views (63 from diffusion, 64 from random cropping and the original image). Table 7 shows the interdependency of λ and $\lambda_\mathbf{y}$ and the role of the *inlierness* scores: as $\lambda_\mathbf{y}$ approaches 0, it tends toward a peak selection and trivial solutions; conversely, as $\lambda_\mathbf{y}$ grows, it tends to MeanShift with uniform *inlierness* scores.

Table 7. Effect of λ and $\lambda_\mathbf{y}$ on the ImageNet dataset. Reported value is the top-1 accuracy averaged over 3 random seeds.

λ \ $\lambda_\mathbf{y}$	0.01	0.05	0.1	0.2	0.4	0.8	1.6	3.2	10	100	$\to\infty$ (MeanShift)
0	66.7	66.7	66.8	68.3	65.8	65.3	65.6	65.9	66.0	66.1	66.1
0.5	66.7	66.7	66.8	68.7	67.7	66.8	66.4	66.2	66.1	66.1	-
1	66.7	66.7	66.9	68.9	68.2	67.4	66.9	66.5	66.3	66.1	-
2	66.8	66.8	67.1	69.1	68.8	68.0	67.4	67.0	66.5	66.1	-
4	66.6	66.5	66.9	**69.3**	69.1	68.6	68.0	67.5	66.8	66.2	-
8	62.0	62.5	64.2	68.7	69.3	69.0	68.5	68.1	67.2	66.3	-
16	57.3	58.5	61.0	65.8	69.1	69.3	68.9	68.5	67.7	66.4	-

Algorithm 1 Mode seeking MeanShift [9]

```
Require: $h > 0$ the bandwidth, $K$ a kernel function (e.g., Gaussian kernel), $\mathbf{m}^0$ a first estimate of the mode, a set of data points $(\mathbf{f}_p)_{1\leq p\leq N}$, a threshold value $\epsilon$
$l \leftarrow 0$
while $l = 0$ or $\|\mathbf{m}^l - \mathbf{m}^{l-1}\| \geq \epsilon$ do
    $\mathbf{m}^{l+1} \leftarrow \frac{\sum_{p=1}^{N} K(\mathbf{f}_p - \mathbf{m}^l)\mathbf{f}_p}{\sum_{p=1}^{N} K(\mathbf{f}_p - \mathbf{m}^l)}$    ▷ mode update
    $l \leftarrow l + 1$
end while
$\mathbf{m} \leftarrow \mathbf{m}^{l-1}$
return $\mathbf{m}$
```

D. Additional results

Zero-shot (Section 5). We report detailed results for Table 1, Table 2 and Table 3 with average top-1 accuracy and standard deviation in Table 8, Table 9 and Table 10 respectively.

Few-shot (Section 6). Additional results for CoOp with 16 tokens are depicted in Figure 5. A similar trend to that shown in Figure 3 is evident, with more pronounced performance degradation observed for TPT. On the contrary, MTA benefits from these more performant prompts.

Ablation study (Section 7). Details for the 15 datasets for the filtering strategy ablation study of Table 6 are given in Table 11. With the exception of ImageNet-A, the confidence threshold strategy consistently demonstrates lower performances compared to our *inlierness* formulation.

Algorithm 2 MTA with Gaussian kernel

Require: A set of augmented embeddings $(\mathbf{f}_p)_{1 \leq p \leq N}$ with $\mathbf{f}_1$ being the original image, a set of class embeddings $(\mathbf{t}_k)_{1 \leq k \leq K}$, a threshold value ϵ, τ the temperature variable of the CLIP model.

```
$w_{p,q} \leftarrow \texttt{Affinity}(\mathbf{f}_p, \mathbf{f}_q, (\mathbf{t}_k)_{1 \leq k \leq K}, \tau) \quad \forall\, p, q \in \{1, ..., N\}$        ▷ See Algorithm 3
$h_p^2 \leftarrow \frac{1}{\rho(N-1)} \sum_{q \in I_p} \|f_p - f_q\|^2 \quad \forall\, p \in \{1, ..., N\}$        ▷ $I_p$ the closest neighbors of p, $\rho$ set to $0.3$
$\mathbf{m} \leftarrow \mathbf{f}_1$        ▷ mode initialization
$y_p \leftarrow \frac{1}{N}\ \forall\, p \in \{1, ..., N\}$        ▷ Initial *inlierness* scores uniform
**while** ***(1)*** and ***(2)*** not converged **do**
    $n \leftarrow 0$
    $\mathbf{y}^0 \leftarrow \mathbf{y}$
    **while** $n = 0$ or $\|\mathbf{y}^n - \mathbf{y}^{n-1}\| \geq \epsilon$ **do**
        $y_p^{(n+1)} \leftarrow \frac{\exp\left((K(\mathbf{f}_p - \mathbf{m}) + \lambda \sum_{q=1}^{N} w_{p,q} y_q^{(n)})/\lambda_{\mathbf{y}}\right)}{\sum_{j=1}^{N} \exp\left((K(\mathbf{f}_j - \mathbf{m}) + \lambda \sum_{q=1}^{N} w_{j,q} y_q^{(n)})/\lambda_{\mathbf{y}}\right)} \quad \forall\, p \in \{1, ..., N\}$        ▷ ***(1)*** *inlierness* scores update
        $n \leftarrow n + 1$
    **end while**
    $\mathbf{y} \leftarrow \mathbf{y}^{n-1}$
    $l \leftarrow 0$
    $\mathbf{m}^0 \leftarrow \mathbf{m}$
    **while** $l = 0$ or $\|\mathbf{m}^l - \mathbf{m}^{l-1}\| \geq \epsilon$ **do**
        $\mathbf{m}^{l+1} \leftarrow \frac{\sum_{p=1}^{N} y_p K(\mathbf{f}_p - \mathbf{m}^l)\mathbf{f}_p}{\sum_{p=1}^{N} y_p K(\mathbf{f}_p - \mathbf{m}^l)}$        ▷ ***(2)*** mode update
        $l \leftarrow l + 1$
    **end while**
    $\mathbf{m} \leftarrow \mathbf{m}^{l-1}$
**end while**
**return** $\arg\max_k\ \mathbf{m}^t \mathbf{t}_k$        ▷ return prediction based on the mode
```

Algorithm 3 Affinity measure based on predictions

```
**function** AFFINITY($\mathbf{f}_p$, $\mathbf{f}_q$, $(\mathbf{t}_k)_{1 \leq k \leq K}$, $\tau$)
    **if** $p = q$ **then**
        **return** $0$
    **end if**
    $l_{p,k} \leftarrow \tau \mathbf{f}_p^t \mathbf{t}_k$ **;** $l_{q,k} \leftarrow \tau \mathbf{f}_q^t \mathbf{t}_k \quad \forall\, k \in \{1, ..., K\}$        ▷ similarity with $\texttt{class}_k$
    $s_{p,k} \leftarrow \frac{\exp l_{p,k}}{\sum_{j=1}^{K} \exp l_{p,j}}$**;** $s_{q,k} \leftarrow \frac{\exp l_{q,k}}{\sum_{j=1}^{K} \exp l_{q,j}} \quad \forall\, k \in \{1, ..., K\}$        ▷ Softmax operation
    $w_{p,q} \leftarrow \mathbf{s}_p^t \mathbf{s}_q$
    **return** $w_{p,q}$
**end function**
```

Table 8. Details of Table 1 with averaged top-1 accuracy and standard deviation computed over 3 random seeds.

Method		**ImageNet**	**-A**	**-V2**	**-R**	**-Sketch**	**Average**
TPT	✗	68.94 ± .06	54.63 ± .21	63.41 ± .12	77.04 ± .02	47.97 ± .05	62.40 ± .03
MTA	✓	69.29 ± .09	57.41 ± .15	63.61 ± .07	76.92 ± .13	48.58 ± .05	63.16 ± .07
MTA + Ensemble	✓	70.08 ± .03	58.06 ± .07	64.24 ± .09	78.33 ± .11	49.61 ± .06	64.06 ± .06
TPT + CoOp	✗	73.61 ± .17	57.85 ± .34	66.69 ± .25	77.99 ± .69	49.59 ± .34	65.14 ± .1
MTA + CoOp	✓	73.99 ± .18	59.29 ± .12	66.97 ± .25	78.2 ± .76	49.96 ± .46	65.68 ± .25

Table 9. Details of Table 2 with averaged top-1 accuracy and standard deviation computed over 3 random seeds.

Method	SUN397	Aircraft	EuroSAT	Cars	Food101	Pets	Flower102	Caltech101	DTD	UCF101	Average
TPT	65.41 ± .03	23.1 ± .39	42.93 ± .2	66.36 ± .31	84.63 ± .03	87.22 ± .19	68.86 ± .32	94.12 ± .21	46.99 ± .31	68.00 ± .22	64.76 ± .05
MTA	64.98 ± 0	25.32 ± .25	38.71 ± .22	68.05 ± .16	84.95 ± .06	88.22 ± .07	68.26 ± .08	94.13 ± .02	45.59 ± .18	68.11 ± .11	64.63 ± .02
MTA + E.	66.67 ± .05	25.2 ± .37	45.36 ± .16	68.47 ± .08	85.00 ± .03	88.24 ± .07	68.06 ± .2	94.21 ± .21	45.9 ± .09	68.69 ± .15	65.58 ± .05

Table 10. Details of Table 3 with averaged top-1 accuracy and standard deviation computed over 3 random seeds.

Augmentation	**Method**	**ImageNet**	**-A**	**-V2**	**R**	**-Sketch**	**Average**
RandomCrop	TPT	68.15 ± .3	51.23 ± .31	66.17 ± .2	76.88 ± .2	49.31 ± .2	62.35 ± .05
	MTA	69.11 ± .4	55.27 ± .15	65.71 ± .4	77.48 ± .36	50.23 ± .4	63.56 ± .11
Diffusion	DiffTPT	67.83 ± .23	53.43 ± .64	65.18 ± .43	76.85 ± .11	50.2 ± .36	62.7 ± .19
	MTA	69.18 ± .4	54.5 ± .31	64.81 ± .1	76.82 ± .26	51.09 ± .4	63.28 ± .07

Table 11. Details of Table 6 for *inlierness* scores ablation study. (1) MeanShift (no *inlierness* scores) (2) confidence thresh. (10%) (3) *Inlierness* scores. I stands for ImageNet, A for ImageNet-A, V for ImageNet-V2, R for ImageNet-R and K for ImageNet-Sketch. Reported values are averaged top-1 accuracy and standard deviation computed over 3 random seeds.

	I	A	V	R	K	SUN397	Aircraft	EuroSAT	Cars	Food101	Pets	Flower102	Caltech101	DTD	UCF101	Average
(1)	66.1 ± .03	48.05 ± .14	60.29 ± .23	67.69 ± .1	40.59 ± .05	63.74 ± .09	25.11 ± .1	24.72 ± .08	66.53 ± .2	83.12 ± .09	85.24 ± .22	66.69 ± .25	91.52 ± .11	44.35 ± .24	65.16 ± .05	59.93 ± .07
(2)	68.26 ± .07	60.66 ± .19	63.3 ± .13	76.14 ± .08	47.59 ± .05	63.56 ± .11	24.52 ± .24	36.13 ± .04	67.59 ± .09	83.39 ± .14	85.83 ± .32	66.51 ± .42	92.69 ± .1	45.45 ± .1	67.41 ± .39	63.27 ± .04
(3)	69.29 ± .09	57.41 ± .15	63.61 ± .07	76.92 ± .13	48.58 ± .05	64.98 ± 0	25.32 ± .25	38.71 ± .22	68.05 ± .16	84.95 ± .06	88.22 ± .07	68.26 ± .08	94.13 ± .02	45.59 ± .18	68.11 ± .11	64.14 ± .01

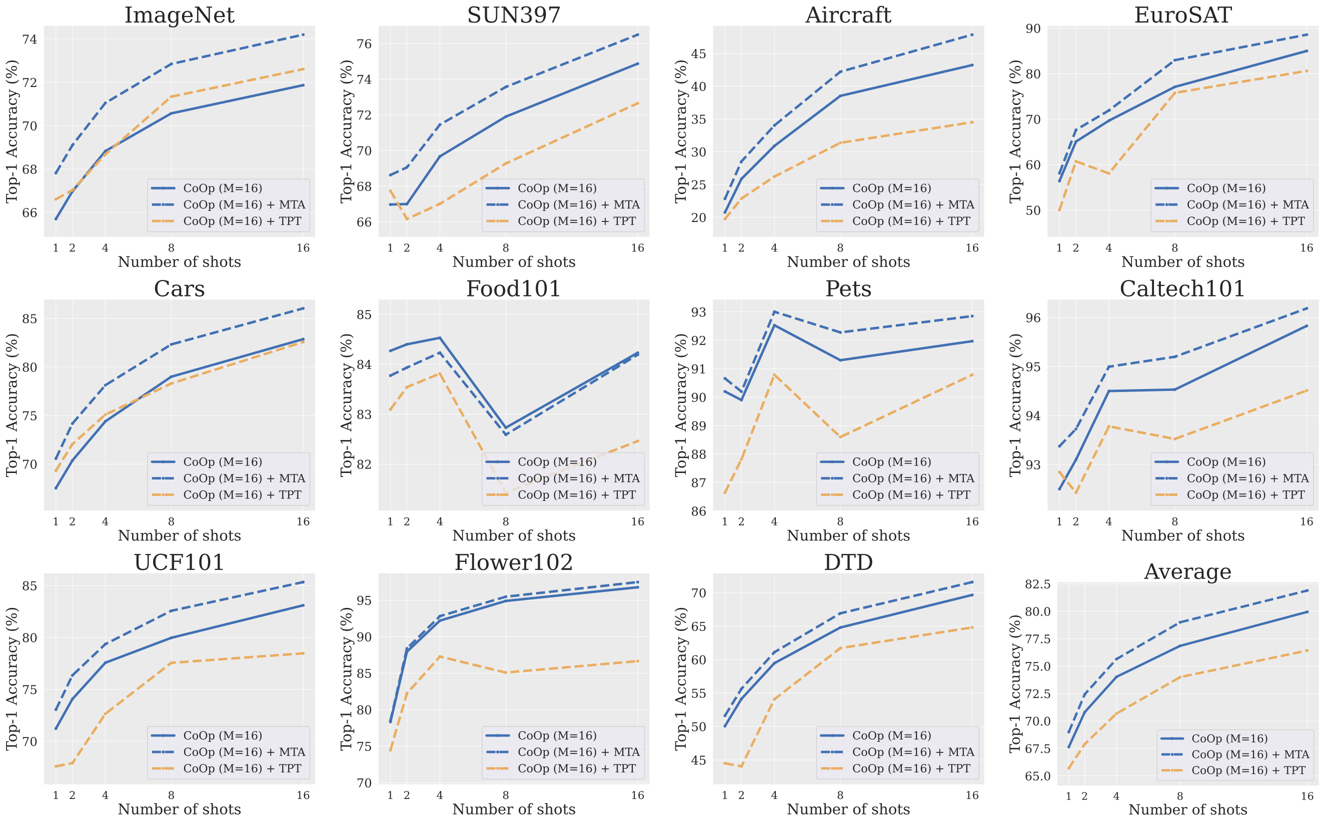

Figure 5. Additional results for Figure 3 with M=16 tokens for the CoOp pretrained prompts.

Table 12. The 80 handcrafted prompts used for majority vote.

”a photo of a [].”, ”a bad photo of a [].”, ”a photo of many [].”, ”a sculpture of a [].”,
”a photo of the hard to see [].”, ”a low resolution photo of the [].”, ”a rendering of a [].”,
”graffiti of a [].”, ”a bad photo of the [].”, ”a cropped photo of the [].”,, ”a tattoo of a [].”,
”the embroidered [].”, ”a photo of a hard to see [].”, ”a bright photo of a [].”,
”a photo of a clean [].”, ”a photo of a dirty [].”, ”a dark photo of the [].”,
”a drawing of a [].”, ”a photo of my [].”, ”the plastic [].”, ”a photo of the cool [].”,
”a close-up photo of a [].”, ”a black and white photo of the [].”, ”a painting of the [].”,
”a painting of a [].”, ”a pixelated photo of the [].”, ”a sculpture of the [].”,
”a bright photo of the [].”, ”a cropped photo of a [].”, ”a plastic [].”,
”a photo of the dirty [].”, ”a jpeg corrupted photo of a [].”, ”a blurry photo of the [].”,
”a photo of the [].”, ”a good photo of the [].”, ”a rendering of the [].”,
”a [] in a video game.”, ”a photo of one [].”, ”a doodle of a [].”,
”a close-up photo of the [].”, ”the origami [].”, ”the [] in a video game.”,
”a sketch of a [].”, ”a doodle of the [].”, ”a origami [].”, ”a low resolution photo of a [].”,
”the toy [].”, ”a rendition of the [].”, ”a photo of the clean [].”, ”a photo of a large [].”,
”a rendition of a [].”, ”a photo of a nice [].”, ”a photo of a weird [].”,
”a blurry photo of a [].”, ”a cartoon [].”, ”art of a [].”, ”a sketch of the [].”,
”a embroidered [].”, ”a pixelated photo of a [].”, ”itap of the [].”,
”a jpeg corrupted photo of the [].”, ”a good photo of a [].”, ”a plushie [].”,
”a photo of the nice [].”, ”a photo of the small [].”, ”a photo of the weird [].”,
”the cartoon [].”, ”art of the [].”, ”a drawing of the [].”, ”a photo of the large [].”,
”a black and white photo of a [].”, ”the plushie [].”, ”a dark photo of a [].”, ”itap of a [].”,
”graffiti of the [].”, ”a toy [].”, ”itap of my [].”, ”a photo of a cool [].”,
”a photo of a small [].”, ”a tattoo of the [].”