Style-Aware Normalized Loss for Improving Arbitrary Style Transfer

Supplementary Material

A. Distribution of VGG-19-based Style Losses

We duplicate Study I in Section 3.2 with VGG-19 [45] as the loss model. Results are presented in Figure 9. As evident, similar conclusion as before can be drawn that the classic style loss does not reflect stylization quality.

B. Derivations

In this section, we present derivations of the upper and lower bounds of the classic AST layerwise style loss (Equation (11) and Equation (12) in the main paper) as discussed in Section 4.2 and restated below.

$$\sup\{\mathcal{L}^l_{\text{AST}_s}(S,P)\}=\frac{\|\mathcal{G}\circ\mathcal{F}^l(S)\|^2+\|\mathcal{G}\circ\mathcal{F}^l(P)\|^2}{N^l}$$

$$\inf\{\mathcal{L}^l_{\text{AST}_s}(S,P)\}=\frac{(\|\mathcal{G}\circ\mathcal{F}^l(S)\|-\|\mathcal{G}\circ\mathcal{F}^l(P)\|)^2}{N^l}$$

Notation. For brevity, we use $\mathcal{G}_S$ to represent $\mathcal{G}\circ\mathcal{F}^l(S)$ and $\mathcal{G}_P$ to represent $\mathcal{G}\circ\mathcal{F}^l(P)$. $\mathcal{G}_{S,k}$ (k=[1, 2, ... N^l]) denotes the kth element in the Gram matrix $\mathcal{G}_S$. We use $\odot$ as the element-wise product between two matrices. N^l is a constant that is equal to the product of spatial dimensions of the feature tensor at layer l.

Proof. The style loss at layer l can be expanded as:

$$\begin{aligned}\mathcal{L}^l_{\text{AST}_s}(S,P) &= \frac{\|\mathcal{G}_S-\mathcal{G}_P\|^2}{N^l}\\ &= \frac{1}{N^l}(\|\mathcal{G}_S\|^2+\|\mathcal{G}_P\|^2-2\times(\mathcal{G}_S\odot\mathcal{G}_P))\end{aligned}$$

Upper bound: We show that term $\mathcal{G}_S\odot\mathcal{G}_P\geq 0$ by writing it into summation format.

$$\mathcal{G}_S\odot\mathcal{G}_P=\sum_{k=1}^{N^l}\mathcal{G}_{S,k}\times\mathcal{G}_{P,k}\geq 0$$

This is because Gram matrix is computed from features output by a ReLU layer, as a result, all values in the matrix are non-negative. The same conclusion can be drawn for other non-negative activation functions, such as sigmoid and softmax, are used. Hence, we have

$$\mathcal{L}^l_{\text{AST}_s}(S,P)\leq\frac{\|\mathcal{G}_S\|^2+\|\mathcal{G}_P\|^2}{N^l}$$

Lower bound: We show that term $\mathcal{G}_S\odot\mathcal{G}_P\leq\|\mathcal{G}_S\|\,\|\mathcal{G}_P\|$ by using Cauchy-Schwarz inequality.

$$\begin{aligned}(\mathcal{G}_S\odot\mathcal{G}_P) &= \sum_{k=1}^{N^l}\mathcal{G}_{S,k}\times\mathcal{G}_{P,k}\\ &\overset{①}{=} \mathcal{G}^{\intercal}_{S,flatten}\cdot\mathcal{G}_{P,flatten}\\ &\overset{②}{\leq} \|\mathcal{G}_S\|\,\|\mathcal{G}_P\|\end{aligned}\tag{14}$$

Step ① is by rewriting summation into a dot product. Step ② is because of Cauchy–Schwarz inequality.

$$\begin{aligned}\mathcal{L}^l_{\text{AST}_s}(S,P) &\geq \frac{1}{N^l}(\|\mathcal{G}_S\|^2+\|\mathcal{G}_P\|^2-2\times\|\mathcal{G}_S\|\,\|\mathcal{G}_P\|)\\ &= \frac{1}{N^l}(\|\mathcal{G}_S\|-\|\mathcal{G}_P\|)^2\end{aligned}$$

□

C. Zoomed In Qualitative Results

We present the zoomed in qualitative results used in Figure 6 in Figures 10 and 11

D. Additional Qualitative Results

We show addtional qualitative results with comparison between the classic AST style loss and our balanced loss in Figures 12 to 15.

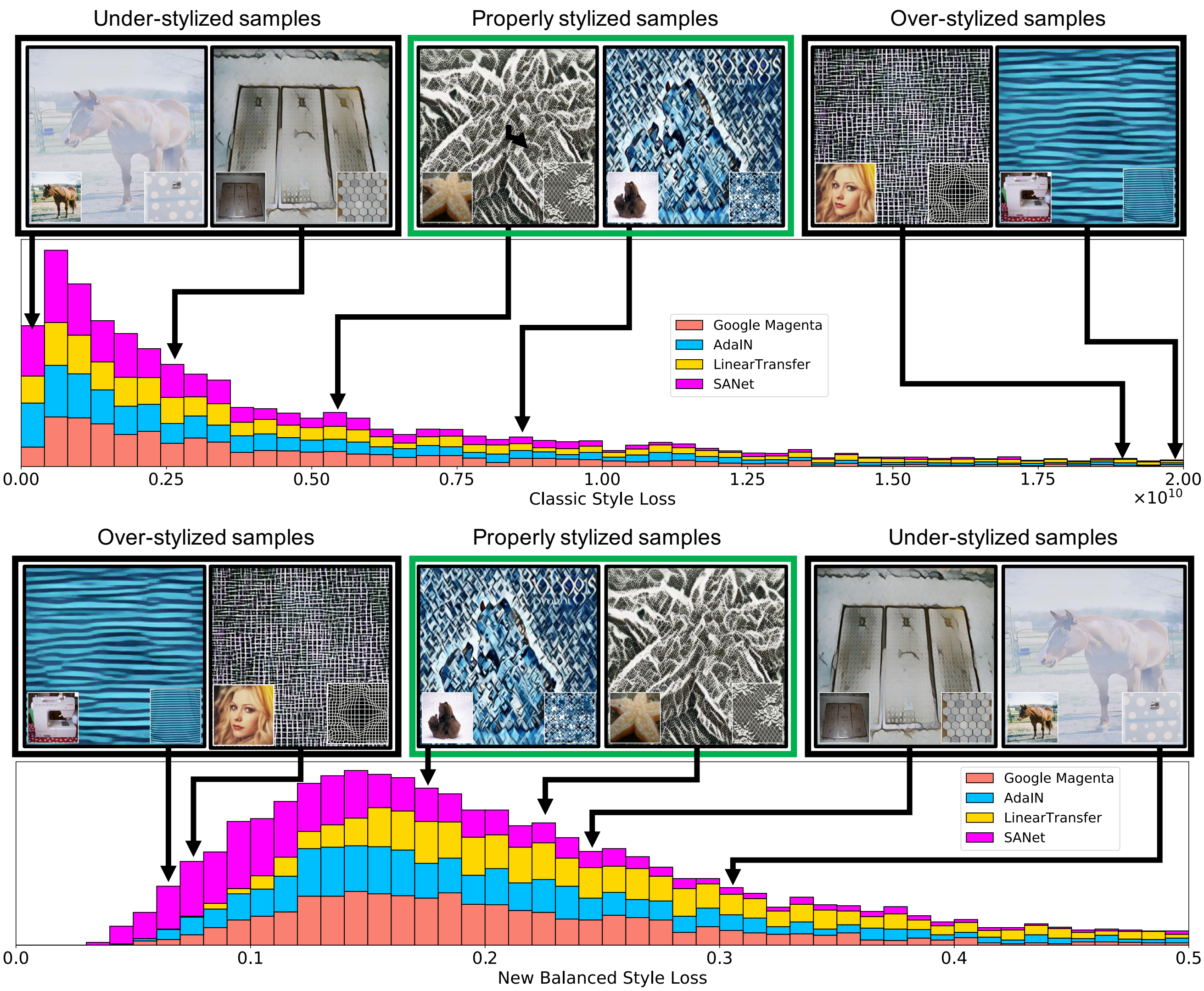

Figure 9: Distribution of the classic Gram matrix-based style losses (above) and balanced losses (below) for the AST methods studied in this work. The distribution of VGG-19-based losses does not reflect stylization quality, same as our previous observation with VGG-16.

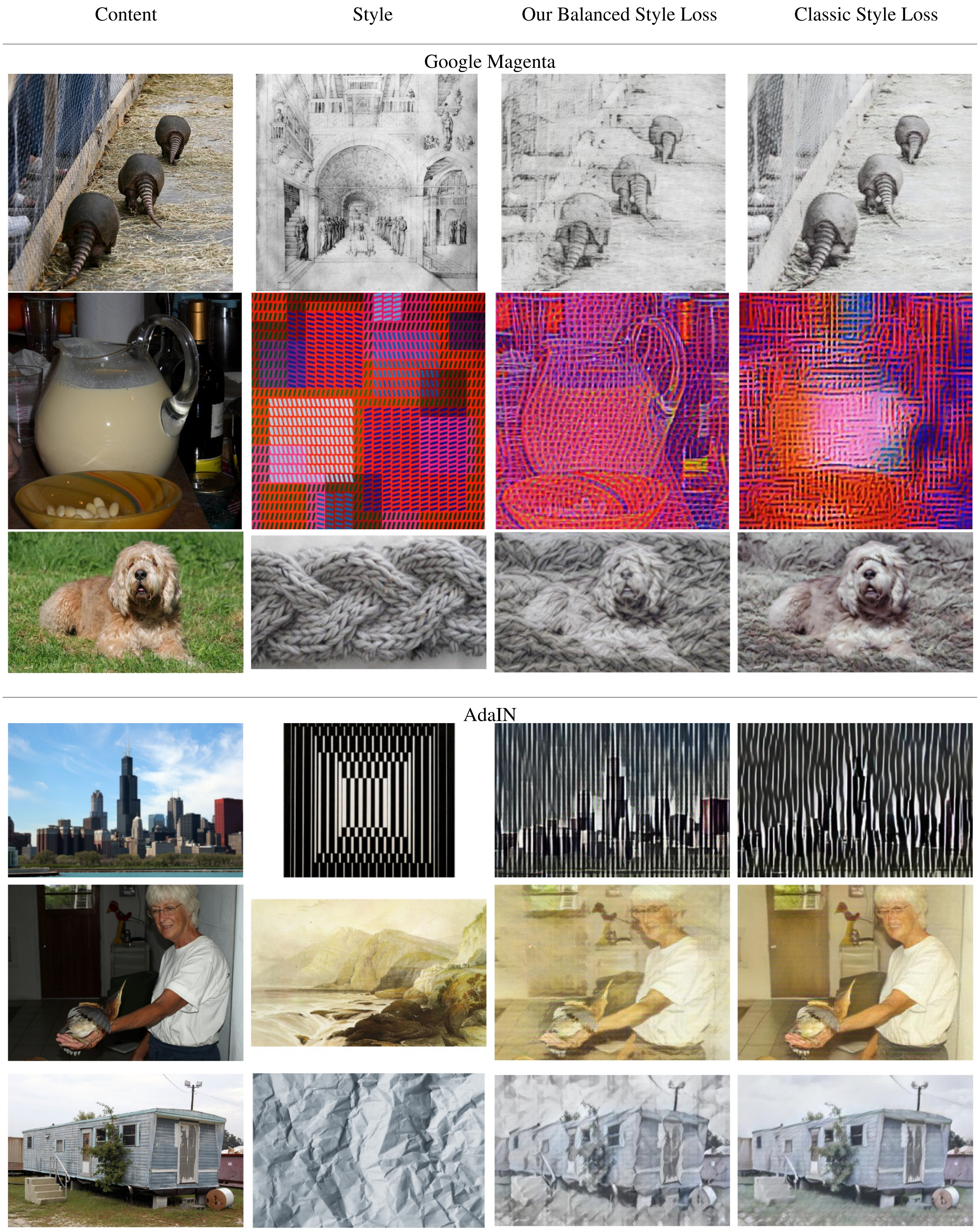

Figure 10: Zoomed in Google Magenta [14] and AdaIN [19] qualitative samples used in Figure 6

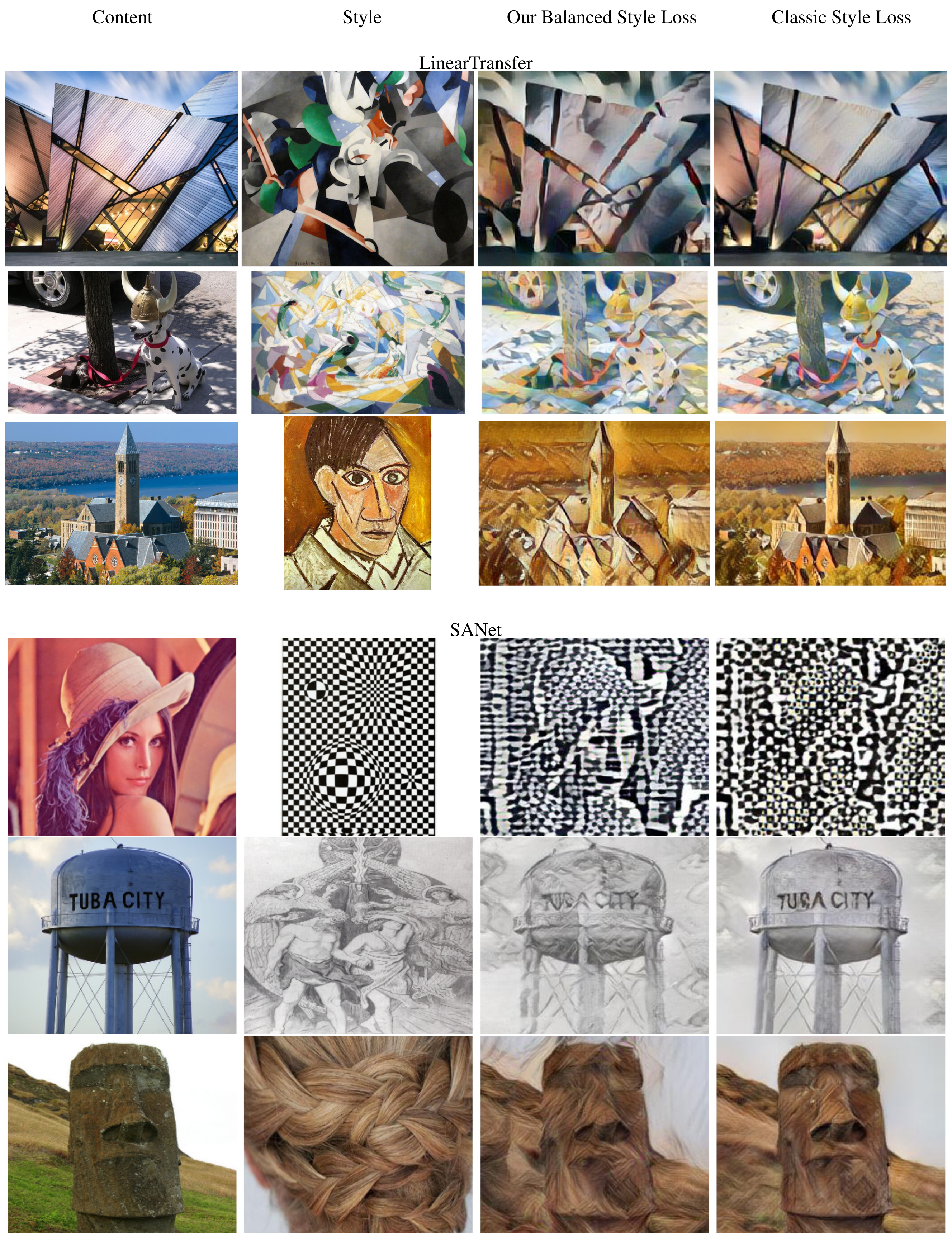

Figure 11: Zoomed in Linear Transfer [29] and SANet [37] qualitative samples used in Figure 6

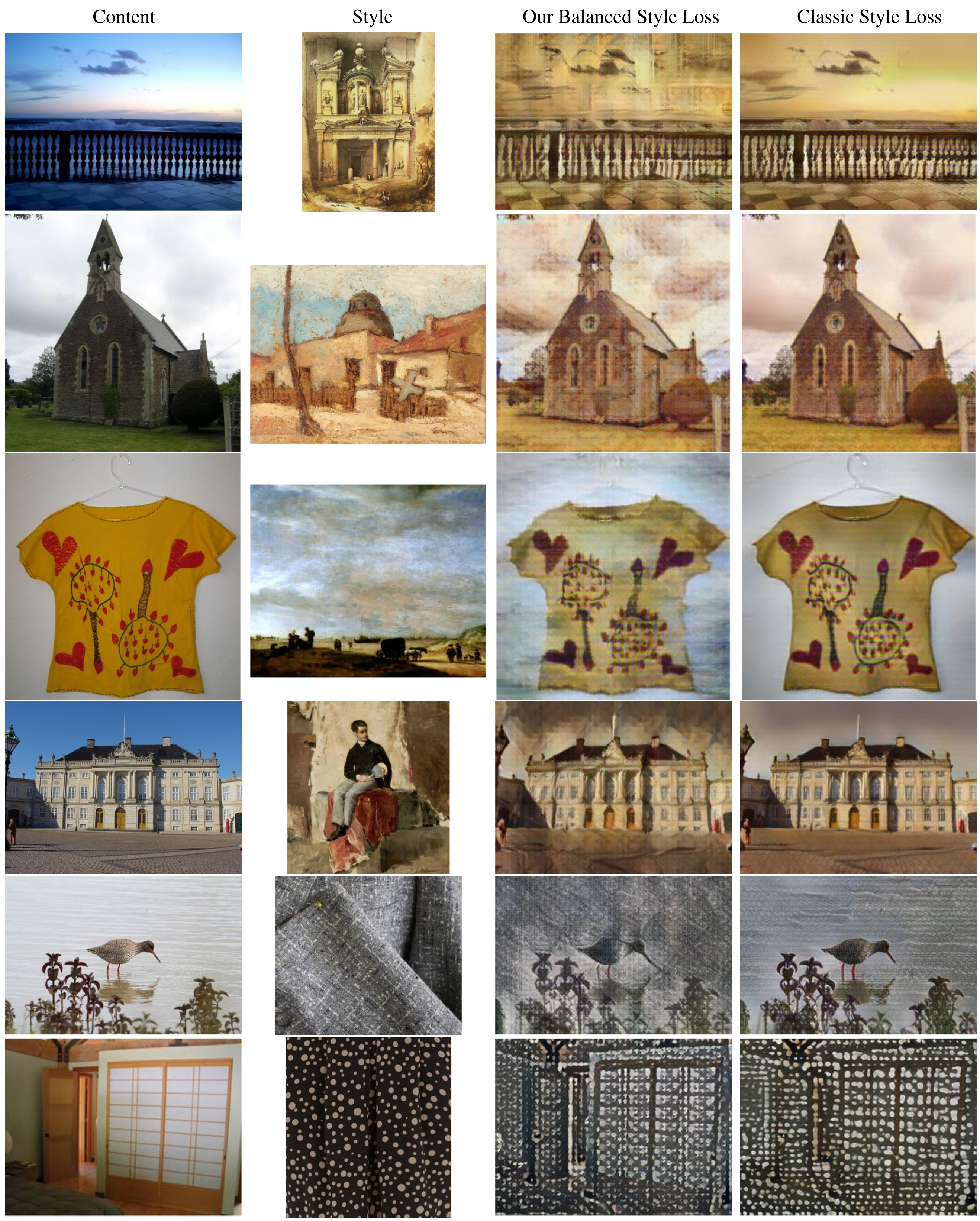

Figure 12: Additional qualitative results for AdaIN [19]

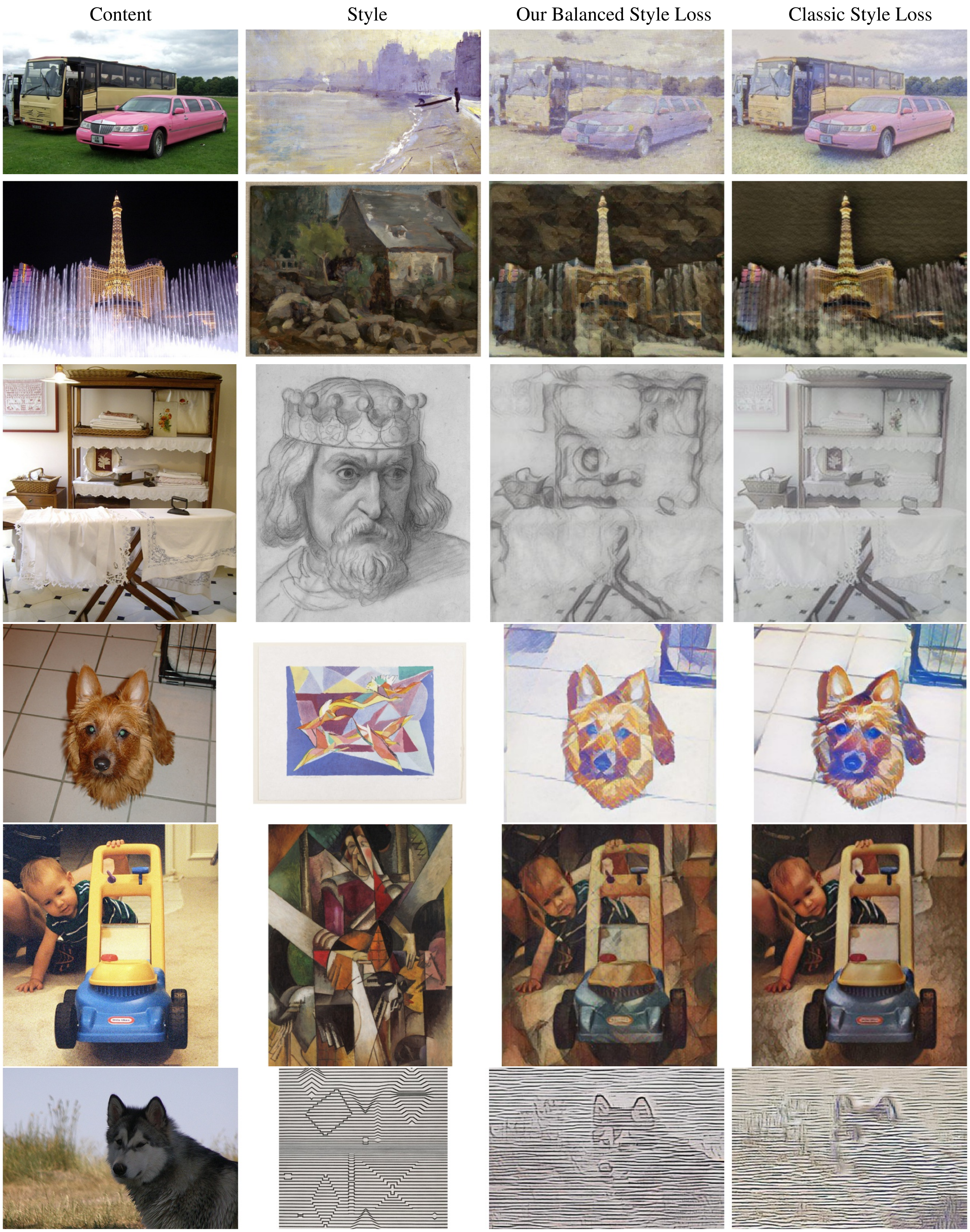

Figure 13: Additional qualitative results for Google Magenta [14]

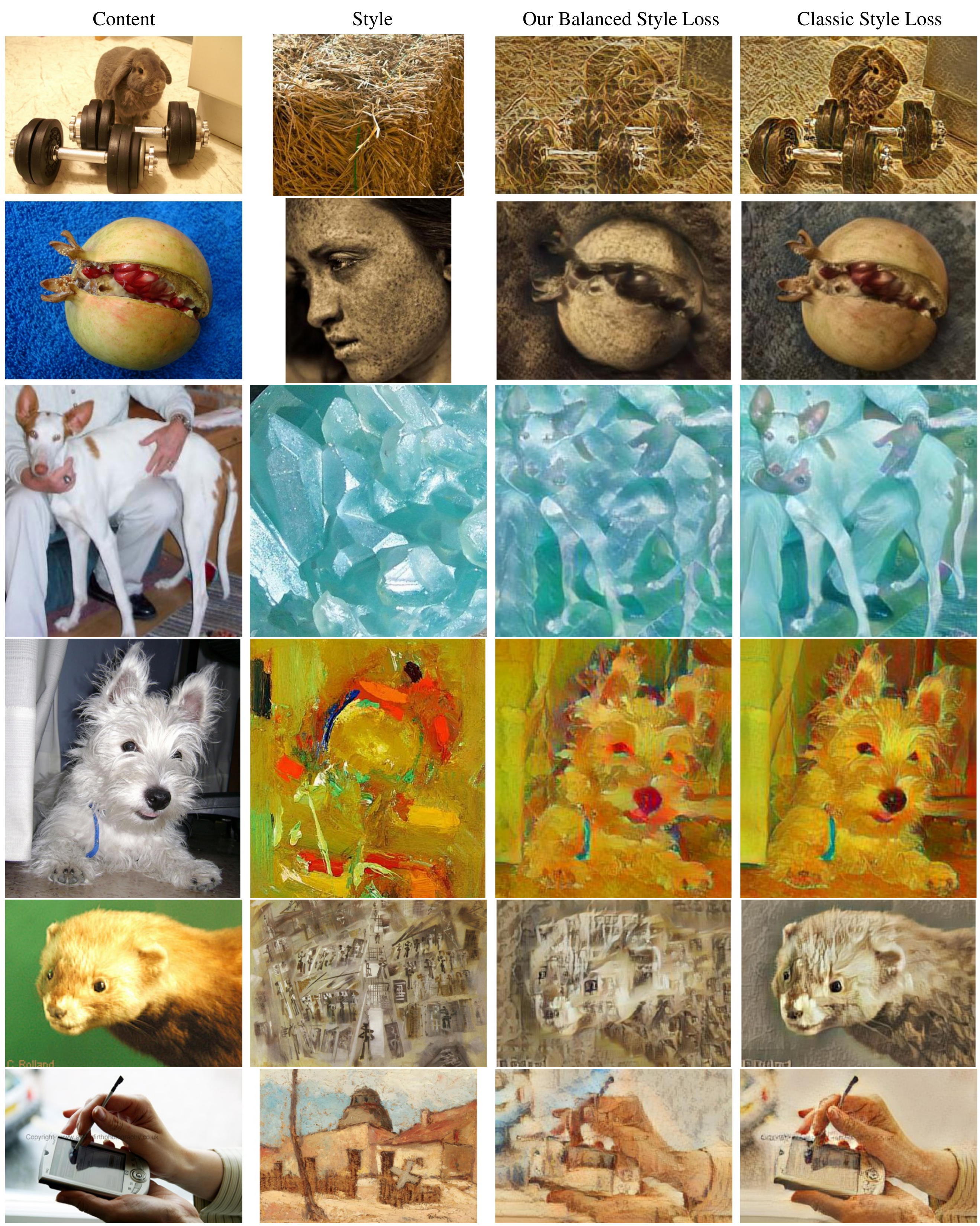

Figure 14: Additional qualitative results for LinearTransfer [29]

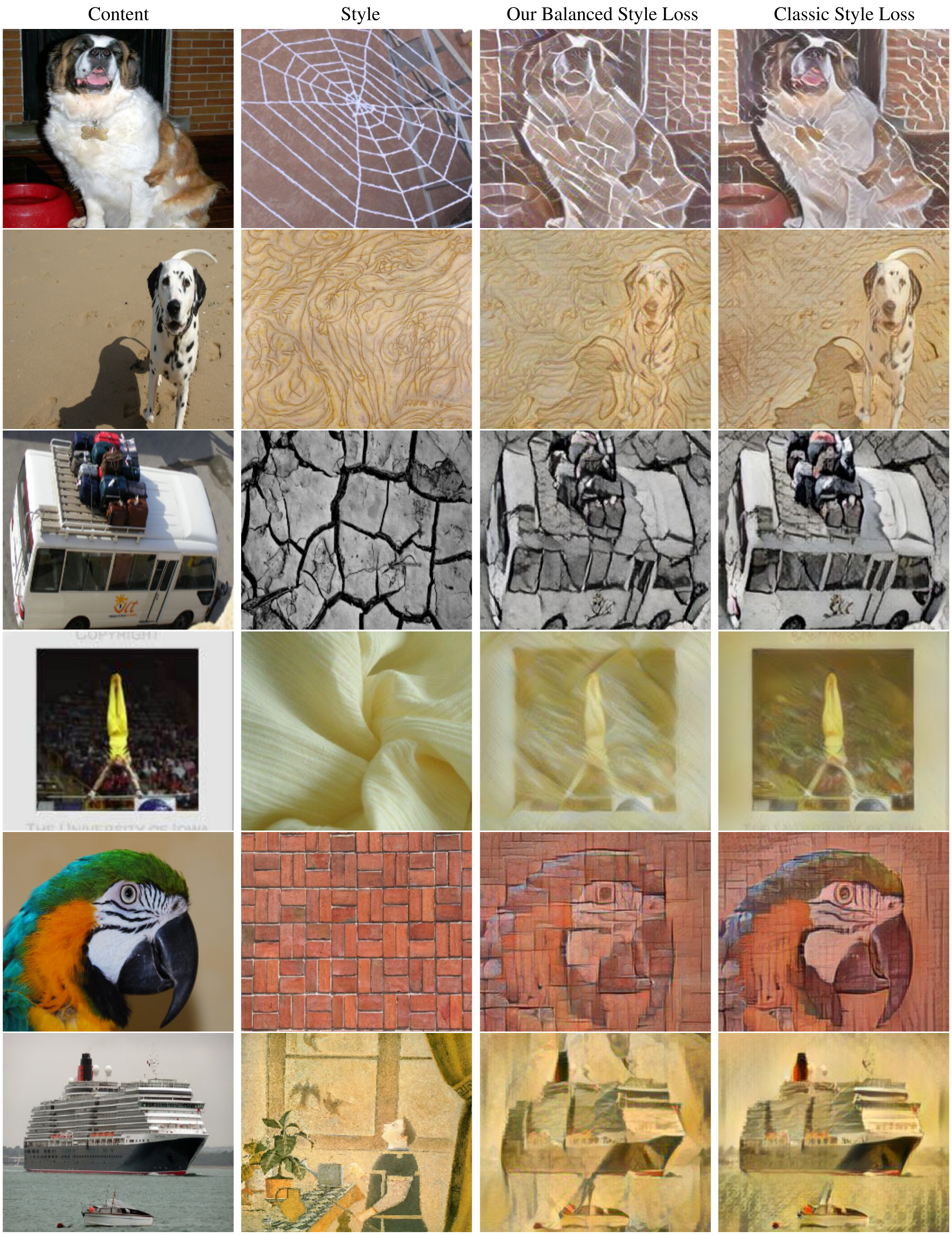

Figure 15: Additional qualitative results for SANet [37]