

**Supplementary Materials
for
STEFANN: Scene Text Editor using Font Adaptive Neural Network**

Prasun Roy^{1*}, Saumik Bhattacharya^{2*}, Subhankar Ghosh^{1*}, and Umapada Pal¹

¹Indian Statistical Institute, Kolkata, India

²Indian Institute of Technology, Kharagpur, India

<https://prasunroy.github.io/stefann>



Figure 1. Editing texts in road signboards with STEFANN. **Left:** Original images. **Right:** Edited images. Text regions are intentionally left unmarked to show the visual coherence of edited texts with original texts without creating passive attention.

*These authors contributed equally to this work.

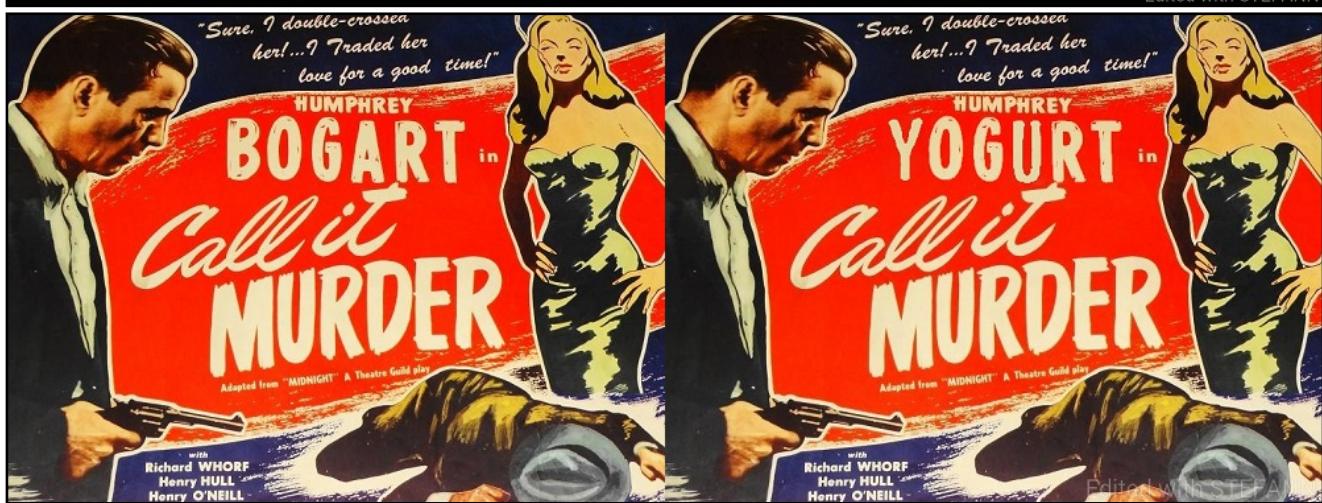


Figure 2. Editing texts in movie posters with STEFANN. **Left:** Original images. **Right:** Edited images. Text regions are intentionally left unmarked to show the visual coherence of edited texts with original texts without creating passive attention.



Figure 3. Editing texts in scene images with STEFANN. **Left:** Original images with text regions marked in red. **Right:** Edited images with text regions marked in green.

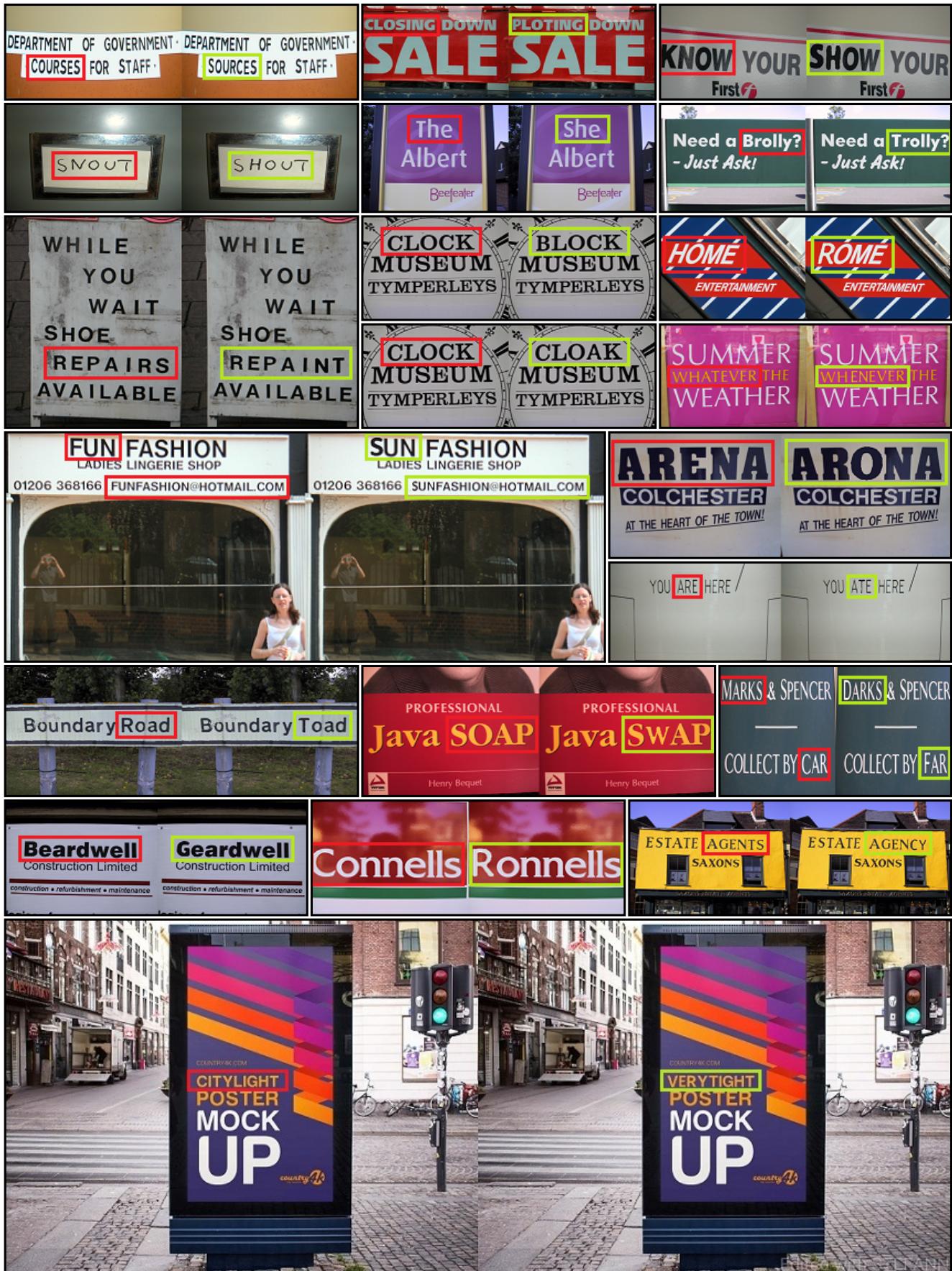


Figure 4. Editing texts in scene images with STEFANN. **Left:** Original images with text regions marked in red. **Right:** Edited images with text regions marked in green.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	w	X	Y	Z
A	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	w	X	Y	Z
B	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	w	X	Y	Z
C	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	w	X	Y	Z
D	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	w	X	Y	Z
E	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	w	X	Y	Z
F	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	w	X	Y	Z
G	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	w	X	Y	Z
H	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	w	X	Y	Z
I	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	w	X	Y	Z
J	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	w	X	Y	Z
K	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	w	X	Y	Z
L	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	w	X	Y	Z
M	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	w	X	Y	Z
N	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	w	X	Y	Z
O	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	w	X	Y	Z
P	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	w	X	Y	Z
Q	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	w	X	Y	Z
R	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	w	X	Y	Z
S	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	w	X	Y	Z
T	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	w	X	Y	Z
U	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	w	X	Y	Z
V	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	w	X	Y	Z
W	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	w	X	Y	Z
X	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	w	X	Y	Z
Y	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	w	X	Y	Z
Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	w	X	Y	Z

Figure 5. Generation of all possible image pairs for a specific font with FANnet. In the first row, characters highlighted in green are the ground truth images. For each subsequent row, character highlighted in red is the image input (source) and characters highlighted in blue are the image outputs (targets) from FANnet, generated by varying the encoding input for each target character. This figure shows the structural consistency of FANnet on a specific font regardless of the source character.

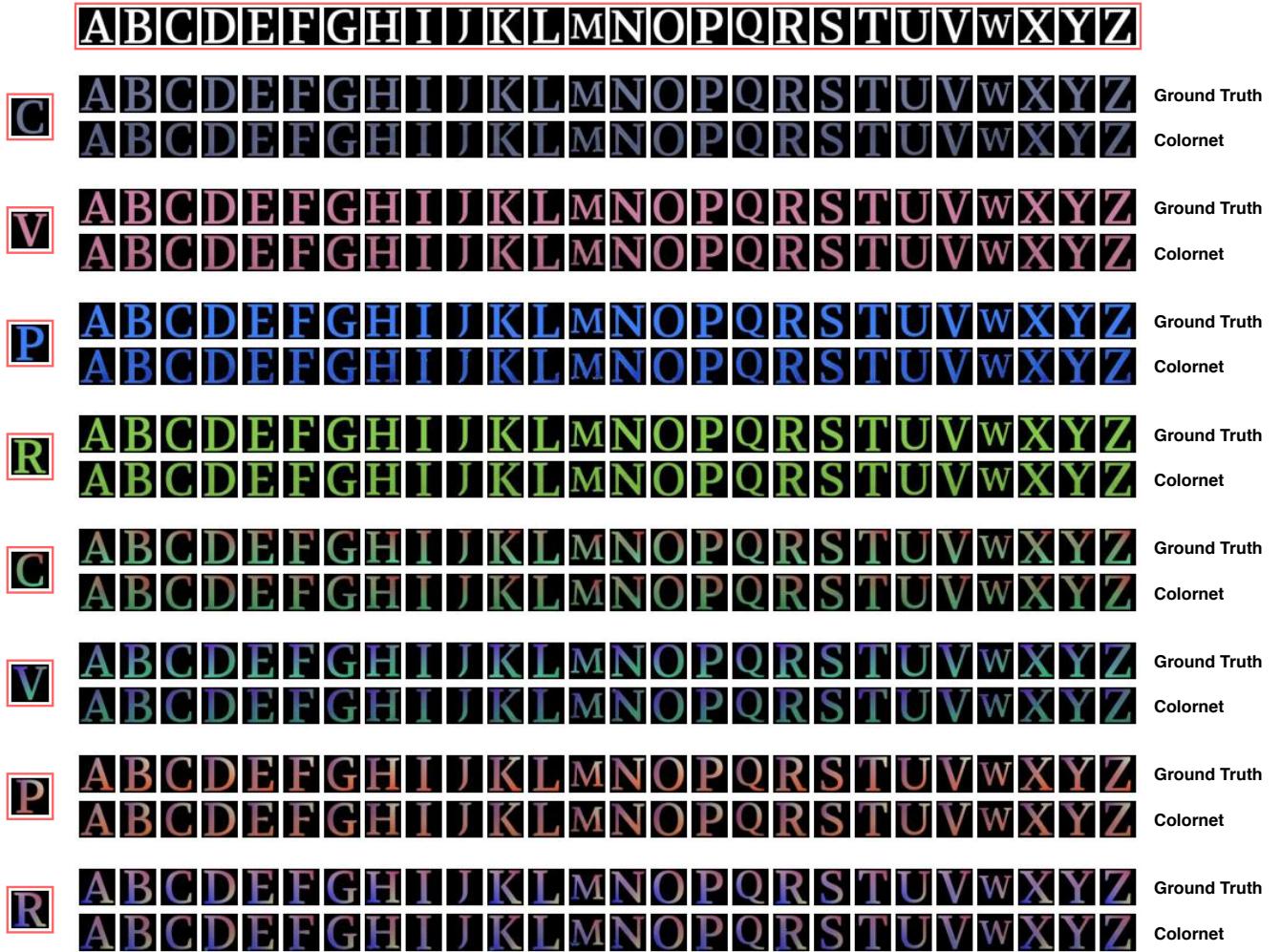


Figure 6. Color transfer with Colornet. In the first row, characters highlighted in red are the binary target character images that need to be colorized. Each subsequent image block shows the colored source character image (highlighted in red), ground truth images for the colored target characters (top) and the color transferred images generated from Colornet (bottom). This figure shows the color consistency of Colornet for both solid and gradient colors.