The Supplemental Material for "DeepVecFont: Synthesizing High-quality Vector Fonts via Dual-modality Learning"

YIZHI WANG, Wangxuan Institute of Computer Technology, Peking University, China
ZHOUHUI LIAN*, Wangxuan Institute of Computer Technology, Peking University, China

Table 1. The detailed configuration of our model. '$k \times k$ $(de)conv$' means the kernel size of a (de)convolutional layer is k. 's' stands for stride of the (de)convolutional layer. 'Out Size' is the size of output feature maps of a block or a (de)convolutional layer (height × width × output channels). 'FC' denotes Fully Connection. 'LN' denotes Layer Normalization [Ba et al. 2016]. 'MDN' denotes Mixture Density Network [Bishop 1994].

Layers	Out Size	Configuration
Image Encoder		
Input	$64 \times 64 \times 1$	-
Layer1	$64 \times 64 \times 16$	7×7 $conv$, s 1×1, LN
Layer2	$32 \times 32 \times 32$	5×5 $conv$, s 2×2, LN
Layer3	$16 \times 16 \times 64$	5×5 $conv$, s 2×2, LN
Layer4	$8 \times 8 \times 128$	5×5 $conv$, s 2×2, LN
Layer5	$4 \times 4 \times 256$	5×5 $conv$, s 2×2, LN
Layer6	$2 \times 2 \times 512$	3×3 $conv$, s 2×2, LN
Layer7	$1 \times 1 \times 1024$	3×3 $conv$, s 2×2, LN
Image Decoder		
Input	$1 \times 1 \times 128$	-
Layer1	$2 \times 2 \times 512$	3×3 $deconv$, s 2×2, LN
Layer2	$4 \times 4 \times 256$	3×3 $deconv$, s 2×2, LN
Layer3	$8 \times 8 \times 128$	5×5 $deconv$, s 2×2, LN
Layer4	$16 \times 16 \times 64$	5×5 $deconv$, s 2×2 , LN
Layer5	$32 \times 32 \times 32$	5×5 $deconv$, s 2×2, LN
Layer6	$64 \times 64 \times 16$	5×5 $deconv$, s 2×2, LN
Layer7	$64 \times 64 \times 1$	7×7 $conv$, s 1×1
Sequence Encoder		
Input	51×10	-
Layer1-4	51×512	LSTM, 512 units, Dropout, LN
Sequence Decoder		
Input	$1 \times 1 \times 128$	-
Layer1-4	51×512	LSTM, 512 units, Dropout, LN
FC	51×4	-
MDN	$51 \times 6 \times 50 \times 3$	-

In this document, we provide additional experimental results in support of the conclusions drawn in our manuscript. The figures to be presented and their corresponding subjects are listed as follows:

- Table. 1 demonstrates the detailed network architecture of our model.
- Fig. 1 and Fig. 2 shows more results of few-shot vector font generation.
- Fig. 3 and Fig. 4 shows the full results of refining from the most similar fonts in our manuscript.
- Fig. 6 shows the full results of vector font interpolation in our manuscript.
- Fig. 7 shows the full results of some randomly generated fonts in our manuscript.

REFERENCES

Jimmy Lei Ba, Jamie Ryan Kiros, and Geoffrey E Hinton. 2016. Layer normalization. *arXiv preprint arXiv:1607.06450* (2016).
Christopher M Bishop. 1994. Mixture density networks. (1994).

---

*Corresponding author

Authors' addresses: Yizhi Wang, Wangxuan Institute of Computer Technology, Peking University, China, wangyizhi@pku.edu.cn; Zhouhui Lian, Wangxuan Institute of Computer Technology, Peking University, China, lianzhouhui@pku.edu.cn.
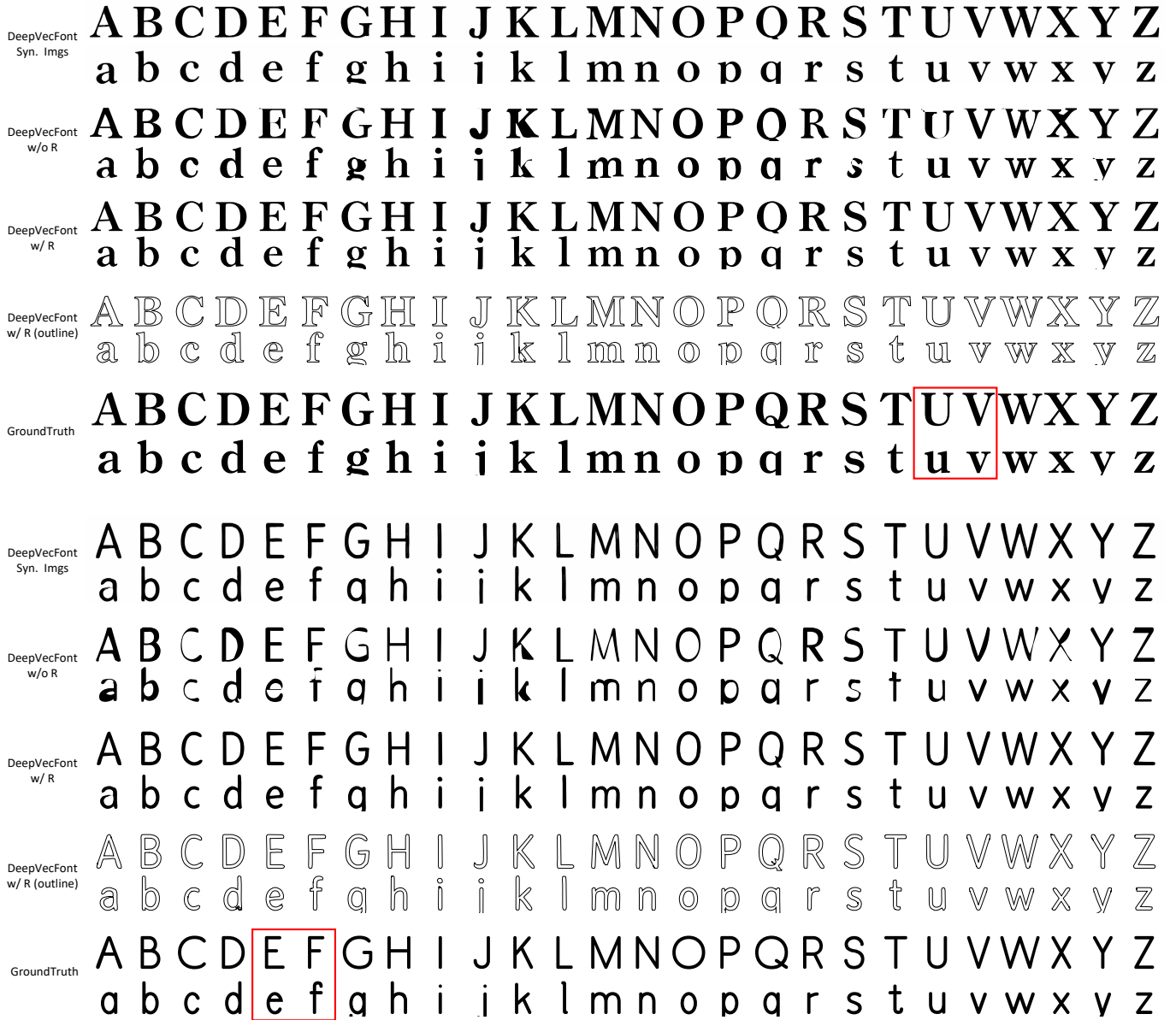
DeepVecFont Syn. Imgs

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 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

DeepVecFont w/o 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
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

DeepVecFont w/ 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
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

DeepVecFont w/ R (outline)

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 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

GroundTruth

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 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

DeepVecFont Syn. Imgs

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 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

DeepVecFont w/o 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
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

DeepVecFont w/ 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
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

DeepVecFont w/ R (outline)

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 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

GroundTruth

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
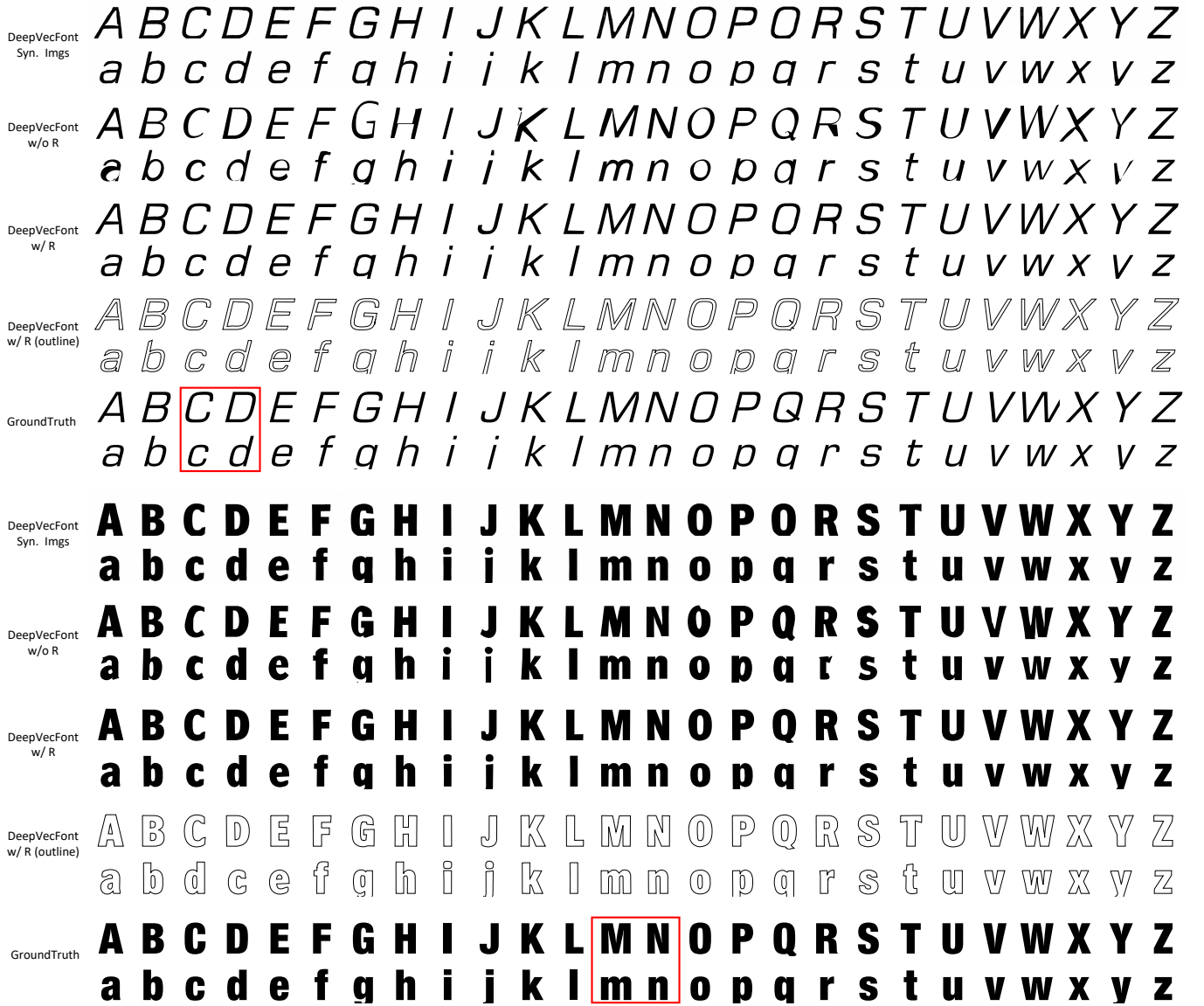a b c d e f g h i j k l m n o p q r s t u v w x y z

Fig. 1. More results of few-shot vector font generation (part-1). The input reference glyphs are marked by red rectangles. "Syn. Imgs" denote synthesized glyph images. "w/o R" denotes the method that synthesizes raw vector glyphs without further refinement.

Fig. 2. More results of few-shot vector font generation (part-2). The input reference glyphs are marked by red rectangles. "Syn. Imgs" denote synthesized glyph images. "w/o R" denotes the method that synthesizes raw vector glyphs without further refinement.

Fig. 3. More results of refining from similar fonts and Arial (part-1). The input reference glyphs are marked by red rectangles. "DVF" denotes DeepVecFont. "Syn. Glyph Imgs" denote synthesized glyph images. "w/o R" and "w/ R" denote without and with refinement according to the synthesized images, respectively.

Fig. 4. More results of refining from similar fonts and Arial (part-2). In this testing font, the lowercase letters' shapes are the same with the upper letters so we only demonstrate the upper ones. The input reference glyphs are marked by red rectangles. "DVF" denotes DeepVecFont. "Syn. Glyph Imgs" denote synthesized glyph images. "w/o R" and "w/ R" denote without and with refinement according to the synthesized images, respectively.

Fig. 5. The outlines of our synthesized glyphs in Fig. 20 of the manuscript. "Ours*" denotes our results presented by outlines. "w/o R" and "w/ R" denote without and with refinement according to the synthesized images, respectively. The input reference glyphs are marked by red rectangles.

Font-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
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

$\lambda = 0.3$

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 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

$\lambda = 0.5$

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 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

$\lambda = 0.7$

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 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

Font-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
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

$\lambda = 0.3$

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 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

$\lambda = 0.5$

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 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

$\lambda = 0.7$

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 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

Font-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
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

Fig. 6. The whole result of vector font interpolations in our manuscript.

Font-1
(Syn. imgs)
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 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

Font-1
(1st sim.)
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 b c d e f g h i j k l m n o p q r s t u v w x u z

Font-1
(Vec.)
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 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

Font-2
(Syn. imgs)
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 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

Font-2
(1st sim.)
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 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

Font-2
(Vec.)
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 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

Font-6
(Syn. imgs)
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 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

Font-6
(1st sim.)
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 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

Font-6
(Vec.)
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
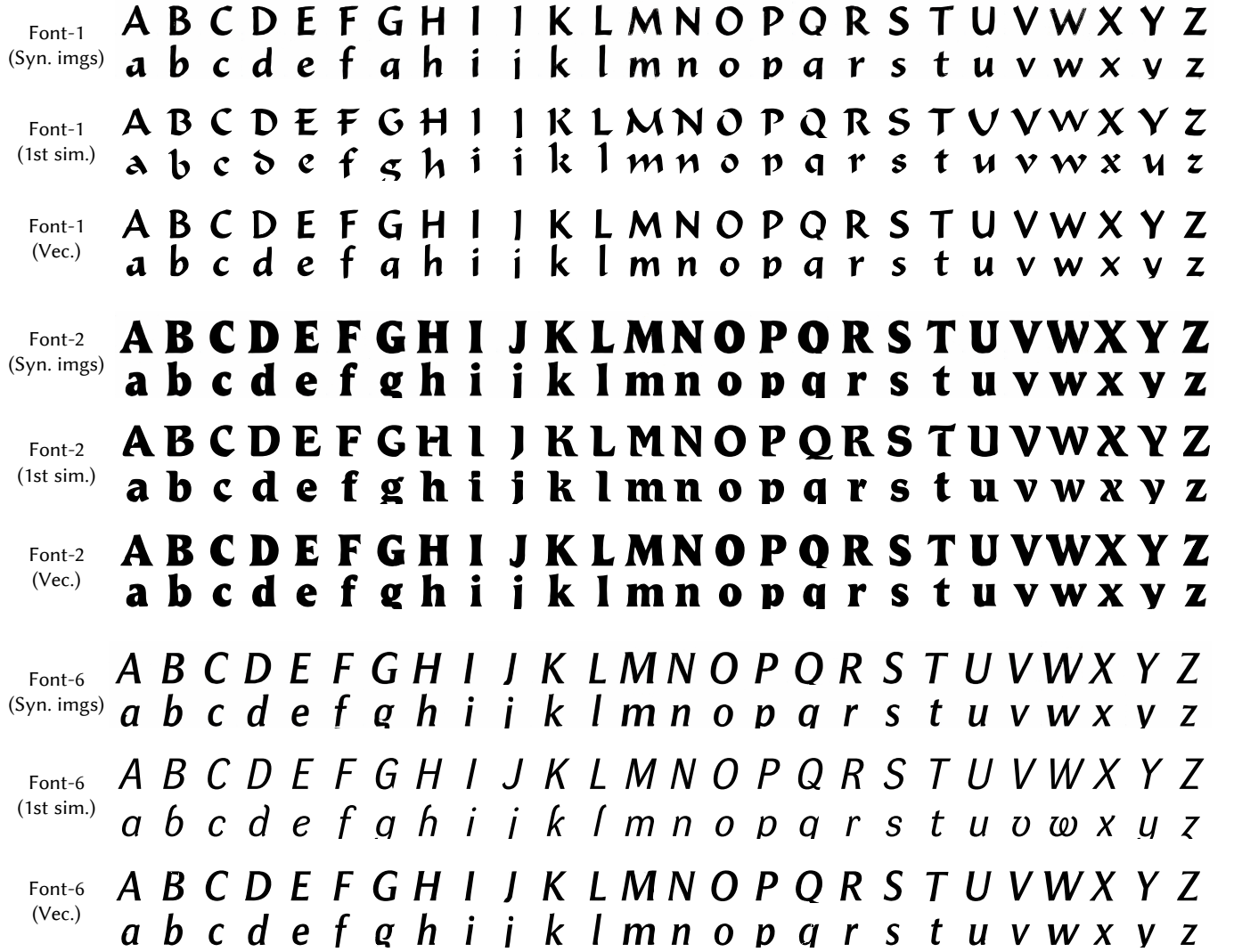a b c d e f g h i j k l m n o p q r s t u v w x y z

Fig. 7. The whole result of randomly generated fonts in our manuscript (Font 1, 2 and 6 in Fig. 19). "Syn. imgs" denotes DeepVecFont's synthesized glyph images. "1st sim" denotes the most similar fonts in the training dataset. "Vec." denotes DeepVecFont's synthesized vector fonts (with refinement).