

Степан Остапенко

Автоматическая генерация наклонных шрифтов

Семестровый проект

Руководитель: Иван Гладких (*TypeType*)

30.05.2023



Факультет математики и компьютерных наук
Программа «Современное программирование»



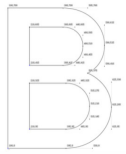



Введение в предметную область

- Компания TypeType занимается разработкой и рисованием цифровых шрифтов.
- Современный шрифт содержит от нескольких сотен до нескольких тысяч символов.
- Каждый символ (глиф) задается в формате SVG (а не JPEG/PNG), т. е. в виде набора команд для его рисования.
- Для каждого шрифта нужно уметь поддерживать разные модификации, например: **жирную** и *наклонную* версии.



Общая постановка решаемой задачи

Разработать механизм для автоматического превращения прямых версий шрифтов в наклонные версии (италики).

	<pre>A: (moveTo', (110, 0,)) (lineTo', (125, 0,)) (lineTo', (1150, 180,)) (lineTo', (505, 180,)) (lineTo', (575, 0,)) (lineTo', (660, 0,)) (lineTo', (410, 700,)) (lineTo', (220, 700,)) (closePath', ()) (moveTo', (235, 280,)) (lineTo', (250, 570,)) (lineTo', (465, 280,)) (closePath', ())</pre>		<pre>A: (moveTo', (20, 0,)) (lineTo', (141, 0,)) (lineTo', (249, 180,)) (lineTo', (559, 180,)) (lineTo', (591, 0,)) (lineTo', (700, 0,)) (lineTo', (569, 700,)) (lineTo', (448, 700,)) (closePath', ()) (moveTo', (211, 280,)) (lineTo', (487, 570,)) (lineTo', (541, 280,)) (closePath', ())</pre>
	<pre>B: (moveTo', (100, 0,)) (lineTo', (505, 0,)) (curveTo', (530, 0, (825, 95), (625, 205)) (curveTo', (825, 330, (505, 375), (505, 375)) (curveTo', (505, 375), (590, 410, (590, 510)) (curveTo', (590, 615, (505, 700), (580, 700)) (lineTo', (100, 700,)) (closePath', ()) (moveTo', (210, 95,)) (lineTo', (210, 325,)) (lineTo', (505, 325,)) (curveTo', (485, 325, (515, 275), (515, 210)) (curveTo', (515, 145, (465, 95), (395, 95)) (closePath', ()) (moveTo', (210, 415,)) (lineTo', (210, 605,)) (lineTo', (580, 605,)) (curveTo', (445, 605, (480, 565), (480, 510)) (curveTo', (480, 455, (440, 415), (380, 415)) (closePath', ())</pre>		<pre>B: (moveTo', (100, 0,)) (curveTo', (100, 0, (416, 0), (426, 0)) (curveTo', (560, 0, (670, 10), (670, 218)) (curveTo', (670, 318, (584, 363), (584, 363)) (curveTo', (584, 363, (700, 398), (700, 530)) (curveTo', (700, 625, (620, 700), (510, 700)) (lineTo', (248, 700,)) (closePath', ()) (moveTo', (230, 95,)) (lineTo', (279, 325,)) (lineTo', (495, 325,)) (curveTo', (530, 325, (561, 280), (561, 222)) (curveTo', (561, 148, (501, 95), (431, 95)) (closePath', ()) (moveTo', (236, 415,)) (lineTo', (338, 605,)) (lineTo', (500, 605,)) (curveTo', (555, 605, (589, 570), (589, 522)) (curveTo', (589, 459, (540, 415), (482, 415)) (closePath', ())</pre>
	<pre>G: (moveTo', (425, -10,)) (curveTo', (820, -10, (725, 125), (725, 125)) (lineTo', (725, 370,)) (lineTo', (470, 370,)) (lineTo', (470, 270,)) (lineTo', (815, 270,)) (lineTo', (815, 165,)) (curveTo', (815, 145, (550, 90), (425, 90)) (curveTo', (280, 90), (165, 205), (165, 350)) (curveTo', (165, 495, (280, 610), (425, 610)) (curveTo', (570, 610, (640, 595), (640, 500)) (lineTo', (710, 575,)) (curveTo', (710, 575, (815, 710), (425, 710)) (curveTo', (425, 710, (55, 550), (55, 300)) (curveTo', (155, 150), (215, -10), (425, -10)) (closePath', ())</pre>		<pre>G: (moveTo', (460, -10,)) (curveTo', (850, -10, (748, 102), (748, 102)) (lineTo', (803, 370,)) (lineTo', (548, 370,)) (lineTo', (527, 270,)) (lineTo', (672, 270,)) (lineTo', (848, 153,)) (curveTo', (848, 153, (574, 90), (467, 90)) (curveTo', (328, 90), (231, 194), (231, 321)) (curveTo', (231, 480), (389, 610), (520, 610)) (curveTo', (660, 610), (727, 498), (727, 498)) (lineTo', (810, 562,)) (curveTo', (810, 561, (734, 710), (527, 710)) (curveTo', (527, 710, (121, 535), (121, 310)) (curveTo', (121, 135, (258, -10), (460, -10)) (closePath', ())</pre>



Исследование и поиск решения

- Поскольку задача является очень сложной, в рамках работы планировалось проведение различных экспериментов и тестирование гипотез.
- Работа основана на статье о предыдущем исследовании в этой области.
- Одним из направлений исследования является проработка двух принципиально разных подходов к решению задачи: растрового и векторного.
- Моя работа посвящена растровому подходу.



Описание подхода к решению

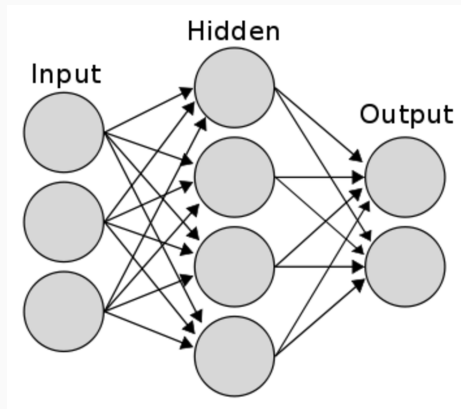
Преобразования глифа: $SVG \rightarrow^1 PNG \rightarrow^2 PNG \rightarrow^3 SVG$

1. растеризатор переводит векторное изображение глифа в пиксельное в заданном разрешении;
2. нейросеть по пиксельному изображению прямого глифа генерирует пиксельное изображение наклонного;
3. кодировщик по пиксельному изображению глифа генерирует набор команд для его рисования.

Многое зависит от правильно выбранного разрешения для промежуточных изображений (PNG).



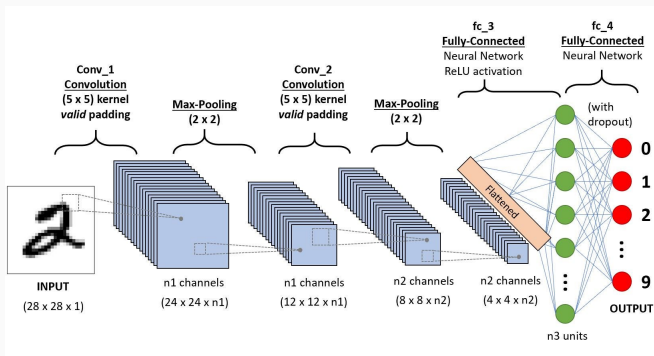
Полносвязная сеть



- На входном слое находятся пиксели изображения прямого глифа, на выходном — наклонного.
- Размеры входного и выходного слоев определяются заданным разрешением (64/128/256), размер скрытого слоя подбирается.



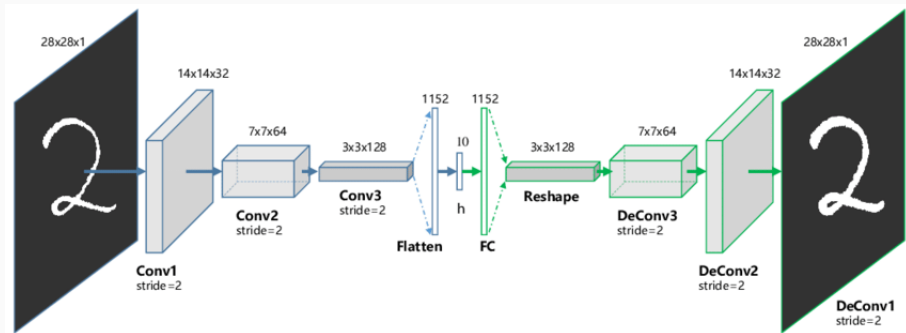
Сверточная сеть



- Основана на архитектуре известной нейросети ResNet18.
- Параметры приходится подбирать отдельно для каждого разрешения (64/128/256).



Генеративная сеть



- Количество параметров не зависит от разрешения, что дает большое преимущество.
- К сожалению, не удалось реализовать в рамках проекта :(8/14



Оценка качества работы

В роли кодировщика использовалась утилита potrace.



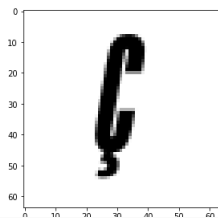
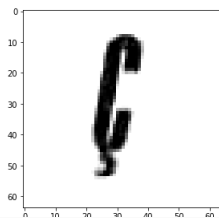
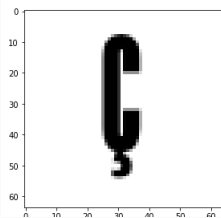
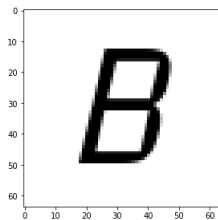
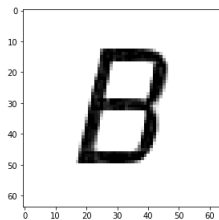
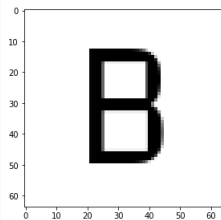
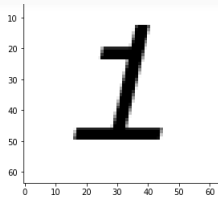
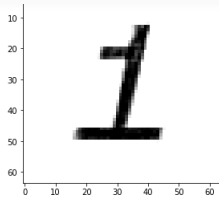
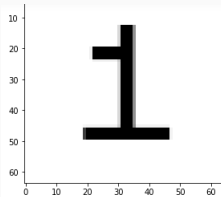
Для оценки результатов работы нейросети считались метрики: MSE, MAE, accuracy, F-score, SSIM. Также учитывалось количество команд для рисования итогового результата.



Метрики и результаты

model	MSE	ACC	FSC	SSIM	mean cmd #
target	0.0	1.0	1.0	1.0	16.125
dense-64	0.007	0.985	0.991	0.928	19.101
dense-128	0.009	0.985	0.991	0.918	35.197
conv-64	0.007	0.920	0.955	0.870	17.751
conv-128	0.008	0.922	0.957	0.826	24.664

- Полносвязная сеть ухудшается с ростом разрешения.
- Сверточная сеть работает стабильнее.
- На выходе кодировщик слишком сильно “смазывает” изображение, из-за чего результаты не подходят.



D

f

á



Выводы по исследованию

- Решение, основанное на растровом подходе очень нестабильное, в первую очередь, из-за кодировщика, который восстанавливает финальные инструкции.
- Для продолжения исследования данного подхода нужно перейти к генеративным моделям для обработки изображений и подобрать настройки для кодировщика (но не факт, что это даст желаемый результат).
- Тем не менее, данный подход можно использовать для получения дополнительной информации о символах, которую можно будет использовать в векторном подходе.



Результаты работы

1. Проведено исследование, получены определенные соображения о дальнейшей работе в этом направлении.
2. Был написан код для предобработки данных, обучения моделей и сравнения результатов, который может быть использован в дальнейших исследованиях.

tg: @flaax, e-mail: stepanos2002@gmail.com

ссылка на репозиторий:

<https://github.com/stephen-ostapenko/auto-italics>

