


MidcurveNN: Encoder-Decoder Neural Network for Computing Midcurve of a Thin Polygon

Yogesh H. Kulkarni 

Yati.io, Pune, India, yogeshkulkarni@yahoo.com

Abstract. Various applications need lower dimensional representation of shapes. Midcurve is one-dimensional(1D) representation of a two-dimensional(2D) planar shape. It is used in applications such as animation, shape matching, retrieval, finite element analysis, etc. Methods available to compute midcurves vary based on the type of the input shape (images, sketches, etc.) and processing (thinning, Medial Axis Transform (MAT), Chordal Axis Transform (CAT), Straight Skeletons, etc.).

This paper talks about a novel method called MidcurveNN which uses Encoder-Decoder neural network for computing midcurve from images of 2D thin polygons in supervised learning manner. This dimension reduction transformation from input 2D thin polygon image to output 1D midcurve image is learnt by the neural network, which can then be used to compute midcurve of an unseen 2D thin polygonal shape.

Keywords: Midcurve, Encoder-Decoder, Neural Network.

INTRODUCTION

A skeleton is a lower dimensional entity which represents shape of its parent object. It being simpler than the parent object, operations like pattern recognition, approximation, similarity estimation, collision detection, animation, matching and deformation can be performed efficiently on it than on the parent object.

Skeletons, also known as Medial Objects, can be computed via various mathematical formulations/approaches such as Medial Axis Transform (MAT), Chordal Axis Transform (CAT), Pairing, Thinning etc. Figure 1 shows some of these. More detailed analysis can be found in the survey paper [1]

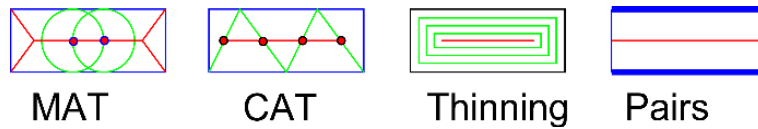


Figure 1: Medial Object computation methods

In the current paper we focus on computing midcurve for 2D planar sketch profiles. Even in 2D profiles, shapes vary enormously. As the first level of simplification, we would deal with 2D polygons only (with an assumption that curved shapes can be converted to polygonal shape by faceting). Figure 2 shows some of the input shapes which can be considered. English alphabets are chosen for easy understanding and verification of the proposed method.



Figure 2: 2D Thin Polygonal shapes

PROPOSED METHOD

Computation of midcurve, in its original form, is transformation of a 2D thin closed, with/without-holes polygon to 1D open/closed/branched polyline. Paper [2] details one of the effective midcurve computation techniques, based on rule-based computational geometry approach. Such techniques have shortcoming of not being scalable or generic enough to be able handle variety of shapes. Deep Learning neural network architectures are showing potential of developing such generic models. This dimension reduction transformation should ideally be modeled as Sequence to Sequence (Seq2Seq) neural architecture, like Machine Translation. In the current problem the input and the output sizes could be different not just in a single sample, but across all samples. Many current Seq2Seq neural networks need fixed size inputs and outputs, which if not present in data, are artificially managed by adding padding of certain improbable value. Such padding is not appropriate for the current midcurve computation problem as the padding-value can inappropriately get treated as part of the valid input. In this initial phase, to circumvent the problem of variable size, image-based inputs and outputs are used, which are of fixed size. Both input and output polygonal shapes are rasterized into images of 100x100, thus making them fixed size for all samples, irrespective of the original shapes.

This paper proposes to use such network for midcurve computation in the form of image-to-image mode. Input images have thin polygonal shapes whereas output images have corresponding midcurve images. Figure 3 shows the Encoder-decoder architecture, called MidcurveNN.

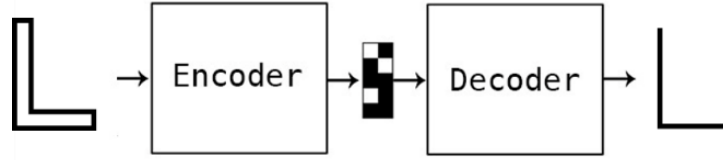


Figure 3: Encoder-Decoder Architecture for Midcurve computation

Input and output geometries are rasterized into 100x100 size images. Transformations like translation, rotation and mirroring are applied to create diversity in the samples. MidcurveNN being a Supervised Learning approach, both input thin-polygons and corresponding output midcurve polylines are transformed simultaneously. Figure 4 shows some samples. This is training data.

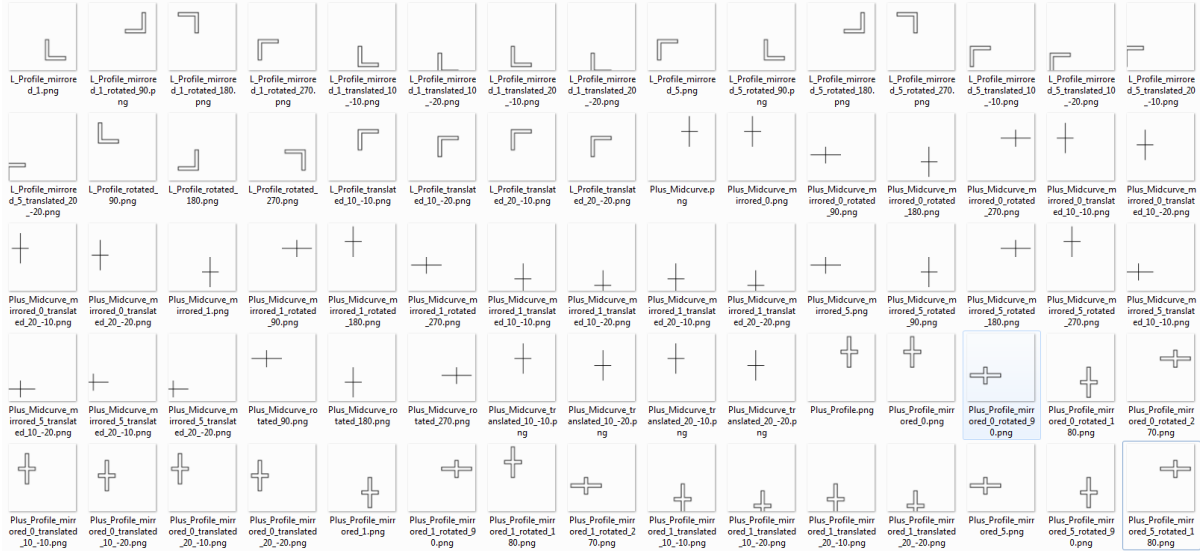


Figure 4: Inputs (thin polygons) and outputs (midcurves)

MidcurveNN encoder-decoder has been implemented in Python programming with Keras library [3]. Encoder takes input image of size $100 \times 100 = 100000$, then comes Dense layer with size 100 to form the encoded vector. Decoder takes encoded vector as input, then with a Dense layer expands back to $100 \times 100 = 10000$ size of the output image. Relu activation is used for Encoder whereas

Sigmoid for the decoder. AdaDelta optimizer with binary cross entropy as loss function is used to compute the losses. Table 1 shows loss across number of epochs.

<i>Epochs</i>	<i>Training loss</i>	<i>Validation loss</i>
50	-17.6354	-8.3223
200	-16.8878	-7.7672

Table 1: Improvement in performance with epochs.

Some of the results are shown in Figure 5. Inputs are in the top and output midcurve at the bottom.

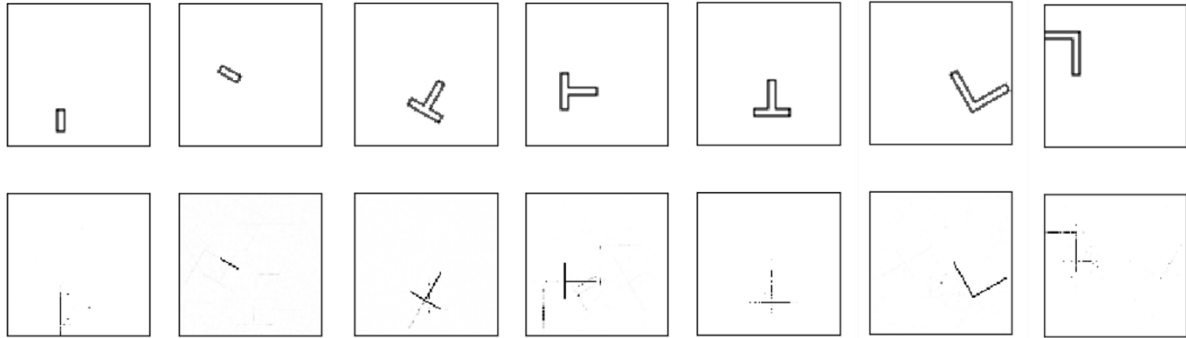


Figure 5: Inputs (thin polygons) and outputs (midcurves)

Shape on the top is the input thin polygon whereas the corresponding shape on the bottom is the predicted midcurve. It can be clearly seen that the network is able to localize the shape and learn the dimension reduction function reasonably well. It is still not perfect or usable, as some stray points are still being wrongly classified as the part of output midcurve. A better network model and/or post processing is needed to make output midcurve practically usable.

CONCLUSIONS AND FUTURE WORK

Traditional methods of computing midcurves are predominantly rules-based and thus, have limitation of not developing a generic model which will accept any input shape. MidcurveNN, a novel Encoder-Decoder network attempts to build such a generic model. This paper demonstrates that simple single layer encoder and decoder network can learn the dimension reduction function reasonably well. Although more development is necessary in evolving a better neural architecture, the current results show positive potential.

Working on truly variable size inputs (thin polygon) and outputs (polyline) using dynamic graph of Encoder-Decoder network can be attempted in the future. More and highly diversified data can help improve the quality of the output. Developing a formal representation of polygonal shapes with variations such as open/closed, with-without loops, branched as a coherent sequence of points is also on the agenda.

REFERENCES

- [1] Kulkarni, Y. H.; Deshpande, S.; Medial Object Extraction - A State of the Art, International Conference on Advances in Mechanical Engineering, SVNIT, Surat, 2010.
- [2] Kulkarni, Y. H.; Sahasrabudhe, A.D.; Kale, M.S.; Dimension-reduction technique for polygons, International Journal of Computer Aided Engineering and Technology, Vol. 9, No. 1, 2017.
- [3] Chollet, F.; Building Autoencoders in Keras, <https://blog.keras.io/building-autoencoders-in-keras.html>