

The Effect of Space-filling Curves on the Efficiency of Hand Gesture Recognition Based on sEMG Signals

Original Scientific Paper

Panagiotis Tsinganos

University of Patras, Department of Electrical and Computer Engineering, 26504 Patras, Greece
Vrije Universiteit Brussel, Department of Electronics and Informatics, 1050 Brussels, Belgium
panagiotis.tsinganos@ece.upatras.gr

Bruno Cornelis

Vrije Universiteit Brussel, Department of Electronics and Informatics, 1050 Brussels, Belgium
IMEC, 3001 Leuven, Belgium
bcorneli@etrovub.be

Jan Cornelis

Vrije Universiteit Brussel, Department of Electronics and Informatics, 1050 Brussels, Belgium
jpcornel@etrovub.be

Bart Jansen

Vrije Universiteit Brussel, Department of Electronics and Informatics, 1050 Brussels, Belgium
IMEC, 3001 Leuven, Belgium
bjansen@etrovub.be

Athanassios Skodras

University of Patras, Department of Electrical and Computer Engineering,
26504 Patras, Greece
skodras@ece.upatras.gr

Abstract – Over the past few years, Deep learning (DL) has revolutionized the field of data analysis. Not only are the algorithmic paradigms changed, but also the performance in various classification and prediction tasks has been significantly improved with respect to the state-of-the-art, especially in the area of computer vision. The progress made in computer vision has produced a spillover in many other domains, such as biomedical engineering. Some recent works are directed towards surface electromyography (sEMG) based hand gesture recognition, often addressed as an image classification problem and solved using tools such as Convolutional Neural Networks (CNN). This paper extends our previous work on the application of the Hilbert space-filling curve for the generation of image representations from multi-electrode sEMG signals, by investigating how the Hilbert curve compares to the Peano- and Z-order space-filling curves. The proposed space-filling mapping methods are evaluated on a variety of network architectures and in some cases yield a classification improvement of at least 3%, when used to structure the inputs before feeding them into the original network architectures.

Keywords – classification, CNN, Deep Learning, electromyography, hand gesture recognition, Hilbert curve, Peano curve, sEMG, space-filling curve, Z-order curve

1. INTRODUCTION

Hand gesture recognition finds many applications in human computer interaction [1], sign language recognition [2], prosthesis control [3] and gaming for rehabilitation [4, 5]. The electrical signals generated by the forearm muscles during the execution of hand gestures provides useful information about muscle activity and hand motion [6]. These signals can be recorded by means of surface electromyography (sEMG) sensors.

The type of hand motions and gestures has tradition-

ally been determined from sEMG data using a combination of (hand-crafted) feature extractors and Machine Learning (ML) classifiers. More precisely, the components of an ML-based pattern recognition system include data acquisition, feature extraction, classification, and inference from new data. In the hand gesture recognition case, the electrodes (sensors) that acquire the sEMG signals, are attached to the arm and/or forearm. Then, features such as Root Mean Square (RMS), variance, zero crossings and frequency coefficients are extracted from the sEMG signals. Finally, these features

are fed into classifiers like k-Nearest Neighbours (k-NN), Support Vector Machine (SVM), Multi-Layer Perceptron (MLP) or Random Forests [7].

The success of Deep Learning (DL) models in solving the problem of sEMG-based hand gesture recognition has been well-documented in numerous studies. In these works, a usual approach is to represent sEMG data as images and feed them into a Convolutional Neural Network (CNN) that outputs a probability that an sEMG signal corresponds a certain gesture class. In a typical CNN architecture, the output label is obtained from a sequence of convolutional and pooling layers followed by fully connected (i.e. dense) layers and a softmax activation layer. Consequently, CNN models transform the input image layer by layer, from its pixel values to the final classification label.

Although an architecture based on Recurrent Neural Networks (RNN) would seem more appropriate for processing time-series signals, little research has been done on this. The problem might be that Long Short-Term Memory (LSTM) networks, the main type of RNN models, require many parameters which cannot be trained efficiently with the limited sEMG datasets available. In addition, recent work [8, 9] has shown that CNN can outperform RNN in a wide range of sequence problems.

There has been considerable progress in the use of CNNs for feature extraction, image classification and particularly for combined feature extraction and classification within the same network. On the other hand, converting time-series into image-like structures that can be used as inputs to CNN models is not a simple task. Methods proposed in the open literature include the partitioning of multi-channel signals using windows and the application of 2D transformations such as the Fourier and Wavelet Transforms.

The current archival periodical article is based on the conference presentation [10], in which we describe a representation method for sEMG signals using the Hilbert curve to obtain images. Now, we investigate the application of several space-filling curves, namely the Hilbert, Peano, and Z-order curves, to represent sEMG signals as images that can be classified by CNNs. Compared to [10], the added contribution presented in this paper is the evaluation of all these types of image representation methods and their performance comparison when applied on the problem of hand gesture recognition using CNNs.

The rest of the paper is structured as follows. Section 2 contains a literature review of the methodologies used for hand gesture recognition. The details of the proposed methods and the CNN architectures applied in our experiments are given in Section 3. The experiments performed for the evaluation of the models are given in Section 4, while the results and a discussion of these results are presented in Section 5. Finally, Section 6 concludes by summarizing the outcomes of this work.

2. RELATED WORK

Hand gesture recognition based on sEMG has been investigated by both typical ML approaches and DL practices. In the case of ML methods, the first approach [11] for the classification of four gestures utilizes time-domain features from two-electrode sEMG signals. An accuracy of 97% in categorizing three types of grasps is achieved in the work of [12] by using the RMS feature of seven electrodes as input to an SVM classifier. The authors of [13, 14, 15] evaluate a large amount of EMG features on a selection of different classifiers for the classification of 52 gestures from the Ninapro reference dataset [13, 16]. The highest performance, an accuracy of 75%, was achieved by a Random Forest classifier using a combination of statistical and frequency domain features.

In the case of DL methods, a large body of research results has been recently developed. The CNN defined in [17] for the recognition of six common gestures resulted in improved accuracy compared to the one obtained by applying an SVM classifier. In [18], a performance comparable to ML approaches is achieved by a model consisting of convolutional and average pooling layers. In [19], we developed techniques to increase the accuracy of the simple model, presented in [18]. Choosing max pooling and inserting dropout layers to reduce overfitting as suggested in [20], allowed us to obtain a 3% increase in accuracy (i.e. from 67% to 70%). The authors of [21, 22] follow a different approach not only in the choice of the model architecture but also in the recording of EMG signals. Their work is based on high-density electrode arrays, which is considered an effective approach in myoelectric control as described in [23, 24, 25]. Using instantaneous EMG images, the CNN model of [21] correctly classifies eight hand movements achieving an accuracy of 89%. The multi-stream CNN described in [22] achieves an accuracy of 85% on the Ninapro dataset, which consists of 52 movements.

Research that utilizes different DL approaches has been carried out as well. The work presented in [26] follows a model adaptation strategy that applies adaptive batch normalization (AdaBN) [27] for updating the parameters of the normalization layers. The use of weighted connections between a source and a target network is presented in [28]. Data augmentation practices for sEMG signals are developed and evaluated in [28]. More recent work in this topic is presented in [29] where magnitude warping and wavelet decomposition showed considerable improvement in classification accuracy.

The use of fractal curves, and in particular the Hilbert curve, for designing alternative image and signal representations is long known, but never studied for sEMG signals. In [30] and [31], the properties of the Hilbert curve have been exploited to convert mammographic images to 1D vectors. In combination with a set of appropriate features, this helped in detecting breast can-

cer. In [32] a similar dimensionality reduction approach is followed. The mapping of 3D data into 2D and 1D representations facilitates the classification of 3D structural data by CNNs. Compared to the direct processing of raw data, such an approach reduces training time and can be used in cases of data of an arbitrary number of channels. The sequence of the extracted image patches is of importance for the detection of image-forgeries by LSTM-based models [33]. The Hilbert curve is employed to determine the order of the image patches fed into the LSTM, thus preserving spatial locality.

The Hilbert curve can be also used in the inverse problem, i.e. transforming 1D data in 2D images. The authors of [34] found that long-term interactions between regions of the DNA sequence are important for its classification. Thus, instead of very deep networks or larger filters, the Hilbert curve was employed to map the DNA sequence into an image such that proximal elements stay close, and the distance between distant elements is reduced.

Besides the more popular Hilbert curve, also Z-order curves have been used in representation problems, though less frequently. For example, in [35] the Z-order curve is chosen because of its good balance between locality preservation and computational complexity to map the neighbourhood around a point in a 3D point cloud into a 1D sequence. These sequences are fed into a CNN to predict the displacement between the current and the next point. The authors of [36] represent the 4D coordinates of a crystal structure as a 1D feature vector using the Z-order curve. An MLP predictor takes as input these representations and estimates the energy of the organic molecular structure.

3. PROPOSED DATA REPRESENTATION METHODS

3.1. Space-filling curves

Space-filling curves, such as the Peano, Hilbert, and Z-order (Morton) curves, have found applications in various domains, such as database access, data compression, and image processing. The key property of

space-filling curves is that they constitute a mapping between a multidimensional space and a lower dimensional space while, in general, they preserve locality between the data points.

The Hilbert curve or Hilbert space-filling curve was first described by the German mathematician David Hilbert in 1891. The main property of this continuous fractal space-filling curve is the superiority in preserving locality compared to alternative curves [37, 38]. The Hilbert curve can be constructed recursively. Firstly, the 2D plane is divided into four quadrants that are traversed according to a fundamental pattern as shown in Fig. 1-(a). In each subsequent iteration, all existing subsquares are subdivided into four smaller subsquares. These four subsquares are connected by a pattern obtained by rotation and/or reflection of the fundamental pattern. Fig. 1-(a) visualizes Hilbert curve traversals of the 2D space after the first iterations, where the numbers are the index within the 1D sequence that is mapped to the specific pixel of the 2D image.

The Peano curve is the first example of a space-filling curve to be discovered, by Giuseppe Peano in 1890. To construct the Peano curve the 2D plane is initially divided into nine squares (i.e. a 3×3 grid) that are traversed according to a fundamental pattern. From each iteration to the next, all existing subsquares are subdivided into nine smaller subsquares connected as shown in Fig. 1-(b).

The Z-order (also known as Morton curve) is another locality preserving space-filling curve named after Guy Macdonald Morton, who first applied the order to file sequencing in 1966. Its basic pattern looks like the letter 'Z'. The curve is constructed as shown in Fig. 1-(c).

3.2. sEMG representation

Fractal curves are used to transform multi-channel sEMG signals into 2D image representations. The sEMG signals are recorded by K electrodes (channels), while performing a hand gesture. These are organized into small segments of N samples. Such an arrangement results in sEMG data of size $N \times K$. The fractal curve map-

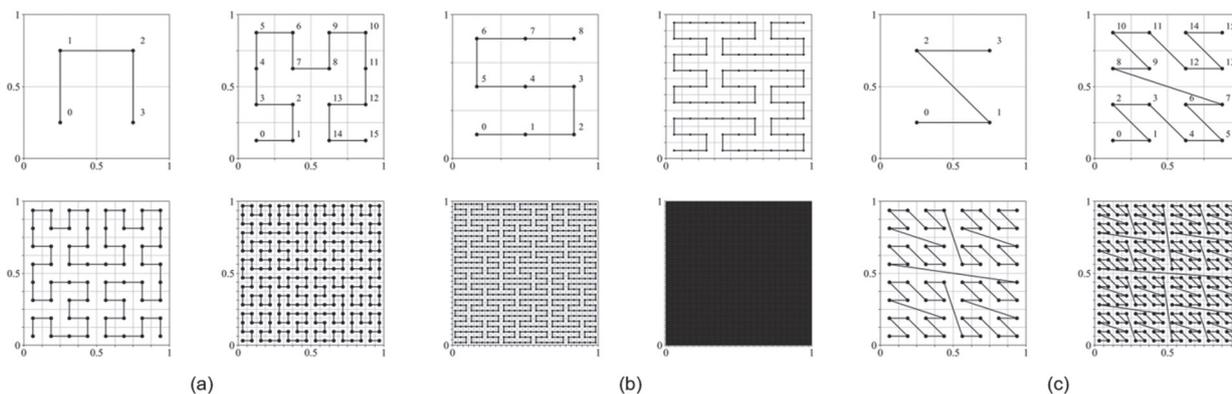


Fig. 1. First iterations of the Hilbert (a), Peano (b) and Z-order (c) curves.

ping can be applied in two ways, either (i) across the time dimension, i.e. map each time sequence of sEMG channel into a 2D image, or (ii) across the sEMG channels, i.e. map the values of the sEMG channels of each time instant into a 2D image (Fig. 2).

The application of the Hilbert and Z-order mappings across the time dimension is as follows. Given a single-electrode sEMG sequence of length N , a 2D representation of $M \times M$ size is produced, where $N=M^2$ and M is a power of two, i.e. $M=2^n$. In the case of K sEMG electrodes, this process is repeated for every electrode, and the outputs are stacked into a K -channel image, resulting in an $M \times M \times K$ image. For example, an sEMG signal of 10 electrodes and 64 samples is mapped into an $8 \times 8 \times 10$ image. The application of the Peano curve is similar, but M should be a power of three, i.e. $M=3^n$. It is important to note that sequence segments of length smaller than M^2 , i.e. $(M-1)^2 < N < M^2$, can be used as well, however, in that case the final image has to be filled out with zeros up to a length of M^2 .

In the second option, the curve mapping is applied across the sEMG electrodes. The number of sEMG electrodes should be a square number, i.e. $K=M^2$, where $M=2^n$ (for the Hilbert and Z-order curves) or $M=3^n$ (for the Peano curve). At every time instant of the sequence the set of K electrode values is mapped into an $M \times M$ image. Thus the dimensions of the Hilbert curve representation of the $N \times K$ segment will be $M \times M \times N$. For example, an sEMG segment of 16 electrodes with 20 samples is mapped into a $4 \times 4 \times 20$ image. As in the previous case, the image will be zero-padded, if there are less electrodes than M^2 (Fig. 2-c).

The computation of the mapping from 1D to 2D using these space-filling curves requires only bitwise operations based on Gray codes performed in $O(1)$ time. For a given $M \times M$ grid, calculating the coordinates for a single data point requires $O(\log_2 M)$ repetitions [39]. Since the mapping is the same for all images, it is computed only once and then used as a look-up table. Therefore, the computational overhead is very small compared to training the CNN.

A common approach in image applications is to use 1D (grayscale) or 3D (RGB) images as inputs to CNNs. In our approach, the dimensionality of the image corresponds either to the number of electrodes or to the sEMG segment duration.

4. EXPERIMENTS

4.1. The dataset

For the evaluation of the proposed models the first dataset of the Ninapro database [13] was selected. This consists of sEMG recordings of 27 healthy subjects that repeat 52 gestures 10 times. The hand movements can be grouped into three categories: (i) basic finger movements, (ii) isometric, isotonic hand configurations and wrist movements, and (iii) grasps and functional movements. EMG signals are measured using 10 electrodes. Eight of them are equally spaced around the forearm. The remaining electrodes are placed on the main activity spots of the large flexor and extensor muscles of the forearm [13].

As in previous studies involving the Ninapro dataset, the sEMG signals are pre-processed by a low-pass

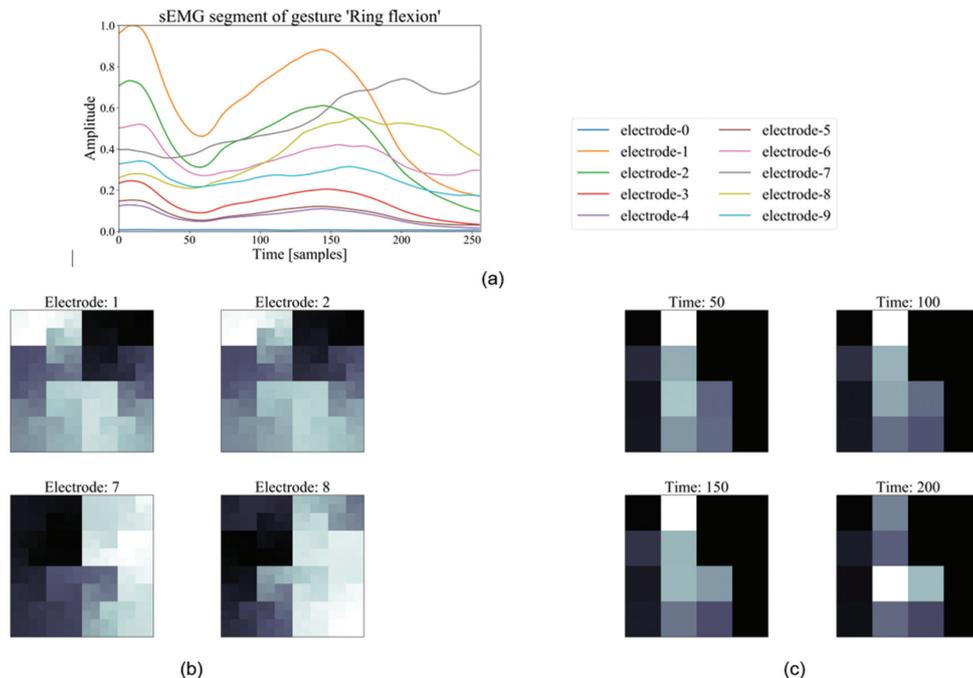


Fig. 2. The application of the Hilbert curve mapping to sEMG data [10]. (a) A 256-samples segment of sEMG signal, (b) the Hilbert representation (16×16) across time of electrodes 1, 2, 6, 8 and (c) the Hilbert representation (4×4) across electrodes at time instants 50, 100, 150, 200. In (c) there are less electrodes than pixel dimensions, thus the last six pixels on the right of the images are zero-filled.

filter [18, 19, 21]. Augmentation of the training data is achieved by duplicating the signals of each repetition and adding Gaussian noise of 25dB signal to noise ratio (SNR). Magnitude-warping is also used for sEMG signal

augmentation [40, 41]. As a last step, sEMG signals from the $K = 10$ channels are segmented into overlapping windows of length N with a step of 50ms and are organized into $N \times 10$ arrays.

Table 1. CNN model architectures used in this work [10]

AtzoriNet [18, 19]	VGGNet [42]	DenseNet [43]	SqueezeNet [44]
CONV(32, 1×10), ReLU,	CONV(16, 3×3), BN, ReLU,	CONV(16, 3×3)	CONV(16, 3×3), BN, ReLU,
CONV(32, 3×3), ReLU,	CONV(32, 3×3), ReLU,	DSBLOCK(4), TRBLOCK(50),	SQBLOCK(4, 16, 32),
POOL(max, 3×3),	CONV(32, 3×3), ReLU,	DSBLOCK(3), TRBLOCK(75),	SQBLOCK(3, 16, 64),
CONV(64, 5×5), ReLU,	POOL(max, 2×2),	ReLU,	CONV(G, 1×1), ReLU,
POOL(max, 3×3),	CONV(64, 3×3), ReLU,	CONV(G, 1×1), ReLU,	GLPOOL(avg),
CONV(64, 5×1), ReLU,	CONV(64, 3×3), ReLU,	GLPOOL(avg),	Softmax
CONV(G, 1×1), Softmax	POOL(max, 2×2),	Softmax	
	GLPOOL(avg),		
	FC(G), Softmax		
		DSBLOCK(n)	SQBLOCK(a, b, c)
		{BN, ReLU, CONV(16, 3×3)} _n	{FIRE(b, c)} _a , POOL(max)
		TRBLOCK(k)	FIRE(b, c)
		BN, CONV(k, 1×1), POOL(max)	CONV(b, 1×1), ReLU,
			CONV(c, 1×1), ReLU + CONV(c, 3×3), ReLU
Parameters	85K	74K	71K
			69K

CONV: convolutional layer, POOL: pooling layer, GLPOOL: global pooling layer, FC: fully-connected layer, BN: batch normalization, { }_k: repetition k times, '+': concatenation
G: number of gesture labels

4.2. Network architectures

A number of neural networks have been used for hand gesture recognition [19]. Apart from those, we also investigate CNN architectures that are typically found in image related tasks, but have not yet been applied to sEMG-based hand gesture recognition, such as VGGNet [42], DenseNet [43], and SqueezeNet [44]. For the comparisons to be fair, an effort is made to keep the number of trainable parameters of the networks approximately equal. The model architectures are shown in Table 1.

4.3. Baseline

As our baseline for comparison, we follow the approach where none of the fractal curve mappings is used. This means that the $N \times 10$ arrays are fed into the CNN models as single-channel images. For the window length N , we experiment with two different values: 16 and 64 samples. The reason for choosing $N = 16$ is that for a real-time application the window size should be as small as possible [11, 45]. We also report results for a bigger window, i.e. $N = 64$, because of the highest accuracy that was achieved during the validation experimentation (Fig. 3). The explanation for the trend in Fig. 3, is that with longer segments ($N > 64$) one can generate less training examples compared to smaller N values. Therefore, the CNN models tend to overfit. On the other hand, too small segments do not contain enough information for the classification task.

4.4. Mapping across time (xxxxTime)

Regarding the space-filling curve mapping across the time dimension (xxxxTime), the $N \times 10$ segments are organized into $M \times M \times 10$ images. In the case of Hilbert and Z-order mappings, for N values equal to 16 and 64 the resulting image sizes are $4 \times 4 \times 10$ and $8 \times 8 \times 10$, respectively, while for the Peano curve the image sizes are $4 \times 6 \times 10$ (zero-padded and cropped) and $9 \times 9 \times 10$ (zero-padded).

4.5. Mapping across electrodes (xxxxElect)

The mapping across the sEMG channel dimension (xxxxElect) is performed in a similar fashion. Given the number of channels $K = 10$, the $N \times 10$ segments are organized into images with dimensions $4 \times 4 \times N$. The pixels corresponding to the last six positions that the three fractal curves traverse are set to zero.

In this approach we retain the spatial resolution constant due to the small number of available electrodes. Regarding the window length, we only experimented with $N = 16$, since a longer segment would generate very deep image representations that in turn would increase the number of parameters in the first convolutional layer increasing the probability of overfitting.

4.6. Model hyper-parameters

All networks were trained using stochastic gradient descent for 60 epochs with an initial learning rate of 0.1

halved every 15 epochs, and a batch size of 1024. To avoid overfitting the networks due to the small training set, dropout layers were appended after each convolutional layer with a forget rate of 0.3. In addition, weight decay regularization with a value of 0.0005 was applied to all convolutional layers. These values were selected after performing a grid search on a validation set of ten randomly selected subjects.

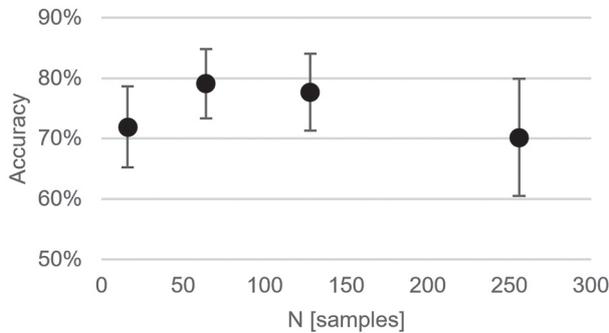


Fig. 3. Validation accuracy vs window size for $N=\{16, 64, 128, 256\}$. Error bars correspond to standard deviation. Highest performance is achieved for $N = 64$.

4.7. Evaluation

The evaluation follows the approach that has been used by other researchers that use the Ninapro dataset [18, 19, 21]. Specifically, a new model is trained for each subject on the data of seven repetitions and is tested on the remaining three. The performance metrics used are the top-1 and top-3 accuracies, i.e. the accuracy when the highest and any of the 3 highest output probabilities match the expected gesture, as well as the precision and recall values. The average across the 27 subjects in the dataset is reported for each metric. For the statistical comparison of the methods, repeated measures Analysis of Variance (ANOVA) and paired Wilcoxon tests are employed, with the F-value being the test statistic and the p-value the corresponding probability.

5. RESULTS

In the experiments, we evaluate two methods for generating sEMG image representations from multi-channel sEMG signals using three different space-filling curves. The evaluation across four CNN models is shown in Table 2, where an ^{‘*’} denotes a significant difference (Wilcoxon

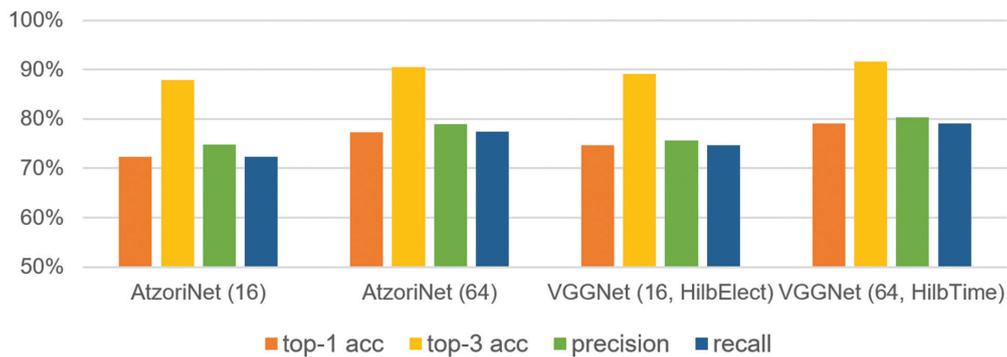


Fig. 4. Comparison of the evaluation metrics for the AtzoriNet[18, 19] and the best performing combinations of representation method and CNN model from Table 2 for $N = 16$ and $N = 64$.

Table 2. Top-1 accuracy metric across sEMG representation method and CNN model. An ^{‘*’} denotes a significant difference ($\alpha = 5\%$) between the baseline and the corresponding space-filling representation. Values in parentheses correspond to the standard deviation.

Segment	Model	Baseline	HilbTime	HilbElect	PeanTime	PeanElect	ZordTime	ZordElect
N = 16	VGGNet	0.7115	0.7192*	0.7469*	0.6957*	0.7346*	0.7045	0.7404*
		(0.0682)	(0.0670)	(0.0653)	(0.0688)	(0.0648)	(0.0685)	(0.0639)
	DenseNet	0.7225	0.7064*	0.7319*	0.6838*	0.7069*	0.6866*	0.7025*
		(0.0639)	(0.0634)	(0.0598)	(0.0689)	(0.0562)	(0.0613)	(0.0597)
	SqueezeNet	0.5611	0.5585	0.5742	0.4062*	0.5934	0.3986*	0.5908
		(0.2342)	(0.1995)	(0.2144)	(0.3017)	(0.1798)	(0.2514)	(0.1816)
N = 64	VGGNet	0.7592	0.7908*	-	0.7527	-	0.7535	-
		(0.0629)	(0.0570)	-	(0.0560)	-	(0.0590)	-
	DenseNet	0.7493	0.7757*	-	0.7544	-	0.7449	-
		(0.0648)	(0.0588)	-	(0.0535)	-	(0.0547)	-
	SqueezeNet	0.6297	0.5188	-	0.3307*	-	0.3281*	-
		(0.1445)	(0.3298)	-	(0.3499)	-	(0.3297)	-

Table 3. Repeated measures ANOVA. On the left, the F and p values of the ANOVA are reported for N = 16 and N = 64. On the right, the p-values of the pairwise comparisons based on the Wilcoxon test are shown. Values above the diagonal correspond to N = 16, and the values below the diagonal are for N = 64. An 'x' denotes an invalid comparison, and an '*' a significant difference ($\alpha = 5\%$).

		HilbTime	PeaNTime	ZordTime	HilbElect	PeaNElect	ZordElect
ANOVA (N = 16)		HilbTime	-	9e-5 *	0.0006 *	x	x
F = 9.6077	p = 8e-8 *	PeaNTime	0.0003 *	-	0.4004	x	x
		ZordTime	0.0010 *	0.6309	-	x	x
ANOVA (N = 64)		HilbElect	x	x	x	-	0.1075
F = 7.7856	p = 0.0011 *	PeaNElect	x	x	x	x	-
		ZordElect	x	x	x	x	-

signed rank test at $\alpha = 5\%$ significance level) between the corresponding representation and the baseline method. Fig. 4 shows the evaluation metrics (top-1 and top-3 accuracies, precision, and recall) for the best performing combination of representation method and CNN model for N = 16 and N = 64. In Table 3, repeated measures ANOVA followed by Wilcoxon signed rank tests assess the significance of the differences between the space-filling curves. On the left, the p-values of the ANOVA are reported, while on the right side of the Table, the pairwise comparisons based on the Wilcoxon test are shown, where the values above the diagonal correspond to p-values for N = 16, and the values below the diagonal are for N = 64. An 'x' denotes an invalid comparison, and an '*' a significant difference ($\alpha = 5\%$).

In general, from Table 2 we see that the Hilbert curve mappings perform always equally well or better compared to the baseline. On the other hand, the image representations of the other two space-filling curves are mostly inferior to the baseline, except for the representations across the electrode dimension (PeaNElect and ZordElect). A comparison between the performances of the CNN architectures, reveals that the VGGNet and the DenseNet yield similar results, whereas the performance of the SqueezeNet is always lower. Considering how much these architectures differ, it is difficult to identify which model components are responsible for this degradation in accuracy performance. The highest top-1 accuracy on the test data is achieved with HilbElect for N = 16 and HilbTime for N = 64, which improves the baseline performance by more than 3%. These representations perform better than the performance of the AtzoriNet architecture, developed specifically for the problem of hand gesture recognition, as shown in Fig. 4.

In Table 3, comparing the top-1 accuracy of the space-filling curves representations shows significant differences for both N = 16 and N = 64. Then, the pairwise comparisons confirm that the sEMG images generated by the Hilbert curve lead to higher performance. In addition, the differences in accuracy between the Peano and the Z-order curves are not significant in general.

For the difference between the representations across time (xxxxTime) and across the electrodes (xxxxElect), we can assume that the former yield better results, since

the dataset used in this work contains data from only ten electrodes. Thus, approximately 1/3 of the pixels of every image correspond to zero padded values, resulting in a limited set of patterns that can be detected by the CNN models compared to the xxxxTime images.

Regarding the advantage of the Hilbert curve in detection accuracy, we can attribute it to the better locality preserving properties between the examined curves. The reason might be that less convolutions are required to extract useful patterns in data, since data are tightly clustered. Therefore, given a network architecture the Hilbert curve image allows for a better utilization of the model parameters.

6. CONCLUSIONS

This work explored the application of the space-filling curves as a means of representing sEMG signals as images for solving the problem of hand gesture recognition. Three curves, i.e. Hilbert, Peano, and Z-order, and two mapping approaches, i.e. traversal across the time dimension and traversal across the electrodes dimension, were evaluated. We made an experimental comparison of three CNN models based on architectures that have been widely used in image processing tasks along with a model optimized for the problem of hand gesture recognition. The results showed that in general, gesture detection accuracy can be improved if sEMG signals are converted to images using fractal curves. The improvement is more evident in the Hilbert curve representations, where the classification accuracy is significantly increased by more than 3% using a VGG-based network. We speculate that this is probably due to the better localization preserving properties of the Hilbert curve. Finally, the differences in accuracy between the Peano and the Z-order curves were not significant in general.

Acknowledgements

The current archival periodical article is based on the conference presentation [10].

The work is supported by the «Andreas Mentzopoulos Scholarships for the University of Patras» and the VUB-UPatras International Joint Research Group on ICT (IJCT).

7. REFERENCES

- [1] S. S. Rautaray, A. Agrawal, "Vision based hand gesture recognition for human computer interaction: a survey", *Artificial Intelligence Review*, Vol. 43, No. 1, 2015, pp. 1–54.
- [2] M. J. Cheok, Z. Omar, M. H. Jaward, "A review of hand gesture and sign language recognition techniques", *International Journal of Machine Learning and Cybernetics*, Vol. 10, No. 1, 2019, pp. 131–153.
- [3] X. Chen, X. Zhang, Z. Y. Zhao, J. H. Yang, V. Lantz, K. Q. Wang, "Hand Gesture Recognition Research Based on Surface EMG Sensors and 2D-accelerometers", *Proceedings of the 11th IEEE International Symposium on Wearable Computers*, Boston, MA, USA, 11-13 October 2007, pp. 11-14.
- [4] Y. J. Chang, S. F. Chen, J. D. Huang, "A Kinect-based system for physical rehabilitation: A pilot study for young adults with motor disabilities", *Research in Developmental Disabilities*, Vol. 32, No. 6, 2011, pp. 2566–2570.
- [5] L. Omelina, B. Jansen, B. Bonnechère, S. Van Sint Jan, J. Cornelis, "Serious games for physical rehabilitation: designing highly configurable and adaptable games", *Proceedings of the 9th International Conference on Disability, Virtual Reality & Associated Technologies*, Laval, France, 10-12 September 2012, pp. 195–201.
- [6] D. Farina, N. Jiang, H. Rehbaum, A. Holobar, B. Graimann, H. Dietl, O. C. Aszmann, "The Extraction of Neural Information from the Surface EMG for the Control of Upper-Limb Prostheses: Emerging Avenues and Challenges", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 22, No. 4, 2014, pp. 797–809.
- [7] E. Scheme, K. Englehart, "Electromyogram pattern recognition for control of powered upper-limb prostheses: State of the art and challenges for clinical use", *The Journal of Rehabilitation Research and Development*, Vol. 48, No. 6, 2011, 643–659.
- [8] S. Bai, J. Z. Kolter, V. Koltun, "An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling", *arXiv:1803.01271*, 2018.
- [9] P. Tsinganos, B. Cornelis, J. Cornelis, B. Jansen, A. Skodras, "Improved Gesture Recognition Based on sEMG Signals and TCN", *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Brighton, UK, 12-17 May 2019, pp. 1169–1173.
- [10] P. Tsinganos, B. Cornelis, J. Cornelis, B. Jansen, A. Skodras, "A Hilbert Curve Based Representation of sEMG Signals for Gesture Recognition", *Proceedings of the 26th International Conference on Systems, Signals and Image Processing*, Osijek, Croatia, 5-7 June 2019, pp. 201–206.
- [11] B. Hudgins, P. Parker, R. N. Scott, "A new strategy for multifunction myoelectric control", *IEEE Transactions on Biomedical Engineering*, Vol. 40, No. 1, 1993, pp. 82–94.
- [12] C. Castellini, A. E. Fiorilla, G. Sandini, "Multi-subject/daily-life activity EMG-based control of mechanical hands", *Journal of NeuroEngineering and Rehabilitation*, Vol. 6, No. 1, 2009, 41.
- [13] M. Atzori, et al, "Electromyography data for non-invasive naturally-controlled robotic hand prostheses", *Scientific Data*, 1, 140053, 2014.
- [14] A. Gijsberts, M. Atzori, C. Castellini, H. Muller, B. Caputo, "Movement Error Rate for Evaluation of Machine Learning Methods for sEMG-Based Hand Movement Classification", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 22, No. 4, 2014, pp. 735–744.
- [15] I. Kuzborskij, A. Gijsberts, B. Caputo, "On the challenge of classifying 52 hand movements from surface electromyography", *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, San Diego, CA, USA, 28 August - 1 September 2012, pp. 4931–4937.
- [16] M. Atzori, et al, "Building the Ninapro database: A resource for the biorobotics community", *Proceedings of the IEEE RAS and EMBS International Conference on Biomedical Robotics and Biomechanics*, Rome, Italy, 24-27 June 2012, pp. 1258–1265.
- [17] K. H. Park, S. W. Lee, "Movement intention decoding based on deep learning for multiuser myoelectric interfaces", *Proceedings of the 4th International Winter Conference on Brain-Computer Interface*, Yongpyong, South Korea, 22-24 February 2016, pp. 1–2.
- [18] M. Atzori, M. Cognolato, H. Müller, "Deep Learning with Convolutional Neural Networks Applied to Electromyography Data: A Resource for the Classification of Movements for Prosthetic Hands", *Frontiers in Neurobotics*, Vol. 10, 2016, pp. 1–10.
- [19] P. Tsinganos, B. Cornelis, J. Cornelis, B. Jansen, A. Skodras, "Deep Learning in EMG-based Gesture Recognition", *Proceedings of the 5th International Conference on Physiological Computing Systems*, Seville, Spain, 19-21 September 2018, pp. 107–114.
- [20] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting", *Journal of Machine Learning Research*, Vol. 15, 2014, pp. 1929–1958.
- [21] W. Geng, Y. Du, W. Jin, W. Wei, Y. Hu, J. Li, "Gesture recognition by instantaneous surface EMG images", *Scientific Reports*, Vol. 6, No. 1, 36571, 2016.

- [22] W. Wei, Y. Wong, Y. Du, Y. Hu, M. Kankanhalli, W. Geng, "A multi-stream convolutional neural network for sEMG-based gesture recognition in muscle-computer interface", *Pattern Recognition Letters*, Vol. 119, 2019, pp.131–138.
- [23] N. Jiang, I. Vujaklija, H. Rehbaum, B. Graitmann, D. Farina, "Is Accurate Mapping of EMG Signals on Kinematics Needed for Precise Online Myoelectric Control?", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 22, No. 3, 2014, pp. 549–558.
- [24] S. Muceli, N. Jiang, D. Farina, "Extracting Signals Robust to Electrode Number and Shift for Online Simultaneous and Proportional Myoelectric Control by Factorization Algorithms", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 22, No. 3, 2014, pp. 623–633.
- [25] A. Stango, F. Negro, D. Farina, "Spatial Correlation of High Density EMG Signals Provides Features Robust to Electrode Number and Shift in Pattern Recognition for Myocontrol", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 23, No. 2, 2015, pp. 189–198.
- [26] Y. Du, W. Jin, W. Wei, Y. Hu, W. Geng, "Surface EMG-Based Inter-Session Gesture Recognition Enhanced by Deep Domain Adaptation", *Sensors*, Vol. 17, No. 3, 458, 2017.
- [27] Y. Li, N. Wang, J. Shi, J. Liu, X. Hou, "Revisiting Batch Normalization For Practical Domain Adaptation", arXiv:1603.04779, 2016.
- [28] U. Côté-Allard, et al, "Deep Learning for Electromyographic Hand Gesture Signal Classification Using Transfer Learning", arXiv:1801.07756, 2018.
- [29] P. Tsinganos, B. Cornelis, J. Cornelis, B. Jansen, A. Skodras, "Data Augmentation of Surface Electromyography for Hand Gesture Recognition," *Sensors*, Vol. 20, No. 17, 4892, 2020.
- [30] S. Dhahbi, W. Barhoumi, J. Kurek, B. Swiderski, M. Kruk, E. Zagrouba, "False-positive reduction in computer-aided mass detection using mammographic texture analysis and classification", *Computer Methods and Programs in Biomedicine*, Vol. 160, 2018, pp. 75–83.
- [31] J. Kurek, B. Swiderski, S. Osowski, M. Kruk, W. Barhoumi, "Deep learning versus classical neural approach to mammogram recognition", *Bulletin of the Polish Academy of Sciences: Technical Sciences*, Vol. 66, No. 6, 2018, pp.831–844.
- [32] T. Corcoran, R. Zamora-Resendiz, X. Liu, S. Crivelli, "A Spatial Mapping Algorithm with Applications in Deep Learning-Based Structure Classification", arXiv:1802.02532, 2018.
- [33] J. H. Bappy, C. Simons, L. Nataraj, B. Manjunath, A. K. Roy-Chowdhury, "Hybrid LSTM and Encoder-Decoder Architecture for Detection of Image Forgeries", *IEEE Transactions on Image Processing*, Vol. 28, No. 7, 2019, pp. 3286–3300.
- [34] B. Yin, M. Balvert, D. Zambrano, A. Schönthuth, S. Bohte, "An image representation based convolutional network for DNA classification", arXiv:1806.04931, 2018.
- [35] A. Thabet, H. Alwassel, B. Ghanem, "MortonNet: Self-Supervised Learning of Local Features in 3D Point Clouds", arXiv:1904.00230, 2019.
- [36] D. Jasrasaria, E.O. Pyzer-Knapp, D. Rappoport, A. Aspuru-Guzik, "Space-Filling Curves as a Novel Crystal Structure Representation for Machine Learning Models", arXiv:1608.05747, 2016.
- [37] C. Gotsman, M. Lindenbaum, "On the metric properties of discrete space-filling curves", *IEEE Transactions on Image Processing*, Vol. 5, No. 5, 1996, pp. 794–797.
- [38] B. Moon, H. V. Jagadish, C. Faloutsos, J. Saltz, "Analysis of the clustering properties of the Hilbert space-filling curve", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 13, No. 1, 2001, pp. 124–141.
- [39] D. Holzmüller, "Efficient Neighbor-Finding on Space-Filling Curves", arXiv:1710.06384, 2017.
- [40] P. Tsinganos, A. Skodras, B. Cornelis, B. Jansen, "Deep Learning in Gesture Recognition Based on sEMG Signals", *Learning Approaches in Signal Processing*, 1st Edition, 2018, p. 471.
- [41] T. T. Um, et al, "Data augmentation of wearable sensor data for parkinson's disease monitoring using convolutional neural networks", *Proceedings of the 19th ACM International Conference on Multimodal Interaction*, Glasgow, UK, 13-17 November 2017, pp. 216–220.
- [42] K. Simonyan, A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition", arXiv:1409.1556, 2014.
- [43] G. Huang, Z. Liu, L. van der Maaten, K. Q. Weinberger, "Densely Connected Convolutional Networks", arXiv:1608.06993, 2016.
- [44] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size", arXiv:1602.07360, 2016.
- [45] C. Attig, N. Rauh T. Franke, J. F. Krems, "System Latency Guidelines Then and Now – Is Zero Latency Really Considered Necessary?", *Lecture Notes in Computer Science*, Springer, Vol. 10276, 2017, pp. 3–14.