Multi-feature Fusion for Epileptic Focus Localization Based on Tensor Representation

Xuyang Zhao*† Jordi Solé-Casals‡§¶ Qibin Zhao† Jianting Cao‖ and Toshihisa Tanaka*†
* Department of Electrical and Electronic Engineering,
Tokyo University of Agriculture and Technology, Japan
† Tensor Learning Team, RIKEN Center for Advanced Intelligence Project, Japan
‡ Data and Signal Processing Research Group, Department of Engineering,
University of Vic–Central University of Catalonia, Catalonia
§ Department of Psychiatry, University of Cambridge, United Kingdom
¶ College of Artificial Intelligence, Nankai University, China
‖ Department of Information Systems, Saitama Institute of Technology, Japan

***Abstract*—Epileptic focus localization based on intracranial electroencephalogram (iEEG) signal is a key task before the patient's surgery, which is time-consuming by visual diagnosis of clinical experts and its accuracy directly affects the effectiveness of surgery. Recently, many machine learning methods have been applied to automatic epileptic focus localization. However, most studies have performed feature extraction and classification based on single type of features, leading to less robustness to noises and degraded performance. In this paper, inspired by making a diagnosis from multiple angles in practical situation, we proposed a multi-set feature fusion strategy by using tensor representation technique. As compared to the existing linear fusion strategy, our proposed tensor fusion approach can significantly enhance the expressive power of the model by taking into account the feature interaction information. We integrate tensor fusion strategy into deep convolutional neural network, yielding a new method for automatic epileptic focus localization. Experimental results demonstrate that our proposed model based on multi-set tensor fusion can achieve the best performance among single type features based models and the linear fusion model.**

I. Introduction

Epilepsy is a common neurological disease of brain caused by the abnormal discharge of brain cells. According to the statistics of the World Health Organization (WHO), approximately 50 million people are suffering from epilepsy and causes a series of troubles in the patient's daily life. Currently, patients with diagnosed epilepsy will be treated with medication first in clinical practice. Then, after long-term drug treatment, part of the patients will be fully recovered, but part of the patients will need to take drugs for whole life. The remaining patients will develop drug resistance causing treatment failure. For these patients, surgical removal of the epileptic focus (lesion) has become an option. Before surgery, clinical experts first needs to locate the focus of the epilepsy. The steps of localization include: (i) the patient's iEEG signal are recorded (usually for one week, and with a minimum of two or three days); (ii) clinical experts evaluate the iEEG signal by visual inspection; and (iii) the final diagnosis is made by a team of experts after discussion and vote. The signal recorded from the epileptogenic area is called focus, and the signal recorded from the non-epileptogenic area is called non-focus, compared to the non-focus signal, the focus signal includes some special waveforms such as spike waves, sharp waves, slow waves, spike-and-slow-waves and so on. Visual assessment is time-consuming, experience-dependent and subjective, the diagnosis by different clinical experts is usually not the same. Because of this, nowadays, the efficiency of localization by physician is low in clinical diagnosis, and patients may even be on hold for treatment. Therefore, it is highly demanding for automatic localization of epileptic focus.

In recent years, many diagnostic aided methods have been proposed that can reduce the workload of clinical experts in visual judgement. Most methods include two steps: (i) feature extraction and (ii) classification. In the feature extract step, several techniques such as empirical mode decomposition (EMD) [1], [2], wavelet transform (WT) [3], [4], [5], entropy [6], [7], [8], phase-amplitude coupling (PAC) [9], [10] are usually used. In the second step, support vector machine (SVM) [11], random forest [12], k-nearest neighbor [13] and neural networks [14] are usually used as a classifier. These methods have helped clinical experts in their diagnosis, but unlike the current diagnosis of clinical experts, these methods often use a single type of feature for training model. However, just use a single type of feature will ignore other hidden nonlinear interaction information and leads to limited expressive power and robustness.

To improve the analyze of iEEG signal more thoroughly, a multi-set feature fusion model is proposed that incorporates efficiently several different types of features. Due to the nonlinear dependence, the non-randomness and the non-stationary characteristics in iEEG signal, the entropy feature is used in the fusion model. As focus signal includes several kinds of abnormal waves such as spike, sharp, slow, spike-and-slow-waves and so on. The abnormal waves have a different frequency distribution. Therefore the frequency feature is extracted by using Short-time Fourier Transform (STFT) and used in the fusion model. In view of the rapid development of convolutional neural networks in recent years, we also use one-dimensional convolutional neural networks to extract features of iEEG signal. In the fusion procedure, through the use of

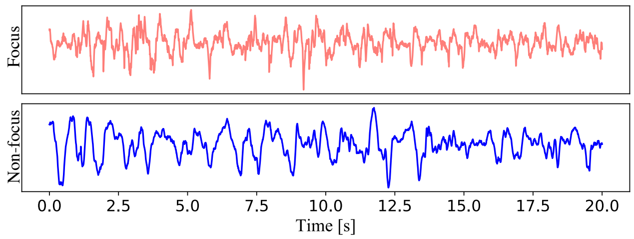

Fig. 1. Samples of focus and non-focus iEEG signal in Bern-Barcelona dataset.

tensor representation, the three features of (i) statistics domain (entropy), (ii) time-frequency domain (STFT) and (iii) graphs (CNN) are merged into one data tensor, which will also allow interactions between features to be taken into account.

The rest of the article is organized as follows: Section 2 introduces the iEEG dataset used for evaluation of the proposed method and describes the three set of features considered and the multi-feature fusion strategy. Section 3 describes the experimental part, while the conclusions are presented in Section 4.

II. Methods

A. Dataset

In this work, the Bern-Barcelona iEEG dataset [15] is used to evaluate the proposed approach. This dataset is recorded by the Department of Information and Communication Technologies of the Universitat Pompeu Fabra (Barcelona, Catalonia) and the Department of Neurology of the University of Bern (Switzerland). Retrospective EEG data analysis has been approved by the ethics committee of the Kanton of Bern. The dataset contains a total of 15,000 samples (7,500 focus samples and 7,500 non-focus samples). The data was collected from five patients suffering from long-standing drug-resistant temporal lobe epilepsy, and candidates for surgery. The signals were recorded with intracranial strip and depth electrodes (AD-TECH, Racine, WI, USA) and preprocessed with a sampling rate of 512 or 1024 Hz, depending on the number of electrodes used (signals sampled at 1024 Hz were first downsampled at 512 Hz, to homogenize the sampling frequency of the dataset). Then, they were band-pass filtered from 0.5 to 150 Hz (fourth-order Butterworth filter). The label to each sample is assigned as follows: the "focus" label was assigned if the channel is in the epileptogenic region; otherwise, the sample is labeled as non-focus. An example of the focus and non-focus iEEG samples are shown in Fig. 1, respectively.

B. Feature Extraction by Using Entropies

Taking into account the nonlinear dependence, the non-randomness and the non-stationary characteristics of the iEEG signal in epilepsy patients, entropy was used as one of the features to extract. The feature extraction was carried out in two steps. Because the different brainwave frequency bands have different physiological meanings, the iEEG signals were first pre-processed using ten three-order Butterworth band-pass filters (0.5–4, 4–8, 8–13, 13–30, 30–50, 50–70, 70–90, 90–110, 110–130, and 130–150 Hz, respectively). The second step was to calculate four types of entropy (Spectral entropy, Approximate entropy, Singular value decomposition entropy, Sample entropy) for each frequency-band signal. Finally, a feature matrix arranged as 11 (raw signal plus the ten frequency-bans) $\times 4$ (entropy measures) was extracted for each sample. The procedure of the feature extraction is illustrated in Fig. 2.

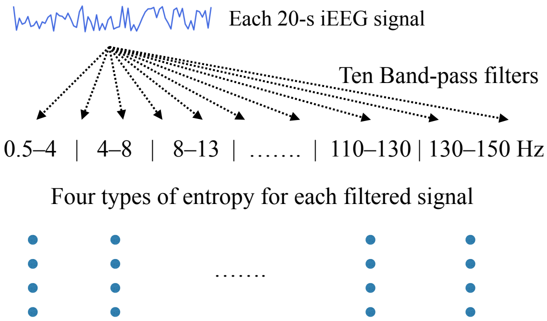

Fig. 2. Each iEEG signal is processed with ten band-pass filters and four types of entropy.

C. Feature Extraction by Using Short-time Fourier Transform

Focus signal includes several kinds of waves such as spike waves, sharp waves, slow waves, spike-and-slow-waves and so on. Each one of them has its own frequency range [16]. Therefore, it will be useful to characterize the iEEG signal from a time-frequency perspective. The Short-time Fourier Transform (STFT) is then used to generate the spectrogram, where the window length is 1 second and the overlap between the windows is 0.8. The output of the STFT feature is a matrix of size 257×101. An example of the spectrogram of the focus and a non-focus sample is shown in Fig. 3. In order to further extract the features in the spectrogram, a shallow 2D-CNN model with the following structure is used: Conv (kernel size = 3, kernel number = 32, strides = 1), Maxpool (pool size = 2, strides = 2) Conv (3, 32, 1), Conv (3, 64, 1), Maxpool (2, 2), Conv (3, 64, 1), Conv (3, 128, 1), Maxpool (2, 2). BatchNorm layer and ReLU function are used behind the Conv layer.

D. One-dimensional Convolutional Neural Network

Recently, the CNN model shows an increasingly important role in multiple fields due to its excellent ability to extract features. CNN has obtained great success in image data processing, where 2D-CNN is normally used. EEGNet [17] is a successful application of 1D-CNN to EEG signal, authors found that using one-dimensional convolution could achieve a filter-like effect. To process the time series data, the 1D-CNN model is used to extract the iEEG wave feature in graphics. The 1D-CNN model architecture is as follows: Conv (kernel size = 3, kernel number = 32, strides = 1), Conv (3, 32, 1), Maxpool (pool size = 4, strides = 4), Conv (3, 64, 1), Conv

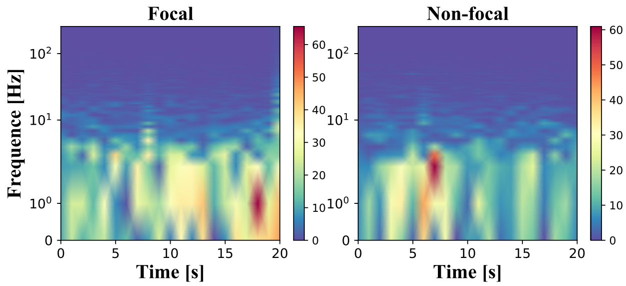

Fig. 3. Examples of spectrogram of focus (left) and non-focus (right) sample, using window length of one second and overlap of 0.8. The X-axis is the time from 0-20 s and the Y-axis is the frequency from 0-256 Hz (in log scale).

(3, 64, 1), Maxpool (4, 4), Conv (3, 128, 1), Conv (3, 128, 1), Maxpool (4, 4). The pooling layer is used to shorten signal length, highlight features, and reduce calculation time, and also can improve spatial invariance to some extent. BatchNorm layer and ReLU function are used behind the Conv layer.

E. Multi-Feature Fusion

In order to simulate clinical experts analyzing iEEG signal from multiple angles, a multi-feature fusion model is proposed, which combines the entropy features, the spectrogram features and the 1D-CNN features at the same time. The overall architecture of our framework is shown in Fig. 4.

When the iEEG signals are fed into the model, three different feature extractions are performed for each sample. After obtaining the three feature vectors (STFT, entropy and 1D-CNN), the vector outer product can be used to represent the data, but this will cause the data to expand rapidly and will cause computational problems. To avoid this issue, a fully connected neural network (FCNN) is added after each feature vector, for the sake of dimensionality reduction. For the STFT and 1D-CNN features, the FCNN includes five layers of size 1024, 512, 256, 128, and 15. For the entropy feature, FCNN includes three layers of size 128, 128, and 15. In all the cases, BatchNorm layer and ReLU function are used behind each layer.

If the fusion is performed by directly calculating the outer product of two feature vectors, the final result is a matrix that contains only cross-feature information; the two original feature vectors are not present in the fused feature matrix. To preserve the original features, an element of ‘1’ is added in each feature vector [18]. Let us consider the case of two features fusion ($\mathbf{z}^1$ and $\mathbf{z}^2$). Using the added element of ‘1’ in the fusion process, the fusion feature $\mathcal{Z}$ can be calculate as:

$$\mathcal{Z} = \begin{bmatrix} \mathbf{z}^1 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} \mathbf{z}^2 \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{z}^1 & \mathbf{z}^1 \otimes \mathbf{z}^2 \\ 1 & \mathbf{z}^2 \end{bmatrix}$$

where $\otimes$ indicates the outer product between vectors. This process is illustrated in Fig. 5, in which four subregions can be identified: the feature vectors $\mathbf{z}^1$ and $\mathbf{z}^2$ (blue and green colours, respectively), the cross feature of $\mathbf{z}^1 \otimes \mathbf{z}^2$ (purple colour), and the constant value ‘1’.

In our model of three features, $\mathbf{z}^e$ is the feature vector of the entropy, $\mathbf{z}^s$ is the feature vector of the spectrogram & 2D-CNN, $\mathbf{z}^c$ is the feature vector of the 1D-CNN. Then, the fusion feature is defined as follows:

$$\mathcal{Z} = \begin{bmatrix} \mathbf{z}^e \\ 1 \end{bmatrix} \otimes \begin{bmatrix} \mathbf{z}^s \\ 1 \end{bmatrix} \otimes \begin{bmatrix} \mathbf{z}^c \\ 1 \end{bmatrix}$$

where $\mathcal{Z} \in \mathbb{R}^{16\times16\times16}$.

After the features are fused with tensor representation, the fused features will be feed into a FCNN for classification. The FCNN model architecture is as follows: Linear(4096, 1024), Linear(1024, 512), Linear(512, 256), Linear(256, 128), Linear(128, 2), BatchNorm layer and ReLU function are used behind the Linear layer (except for the last layer).

III. Experimental results and discussion

Using the Bern-Barcelona dataset, five classification models are evaluated implementing a 10-fold cross-validation strategy. The model architectures are listed in Table I. The following models are used, which differ from each other depending on the type of features:

- For the entropy features, an SVM model with the kernel of radial basis function (RBF). The hyper-parameters ’C’ and ’gamma’ are adjust by grid search (C: [1e-1, 1e1, 1e2, 1e3, 1e4], gamma: [1e-2, 1e-1, 1e1, 1e2, 1e3]).
- For the STFT features, a FCNN is used as a classifier, which includes five layers of size 1024, 512, 256, 128, and 2. BatchNorm layer and ReLU function are used behind each layer.
- For the 1D-CNN features, the same FCNN used in the STFT case is adopted as the classifier.
- For the linear fusion model, the three sets of features are fused by linear concatenation, obtaining a global feature of size of 1×45. Then, the fused features are fed into a FCNN used as a classifier, which includes three layers of size 128, 128, and 2. BatchNorm layer and ReLU function are used behind each layer.
- For the tensor fusion model, three sets of features are fused using the tensor fusion procedure, obtaining a tensor of $16 \times 16 \times 16$. Then, the tensor features are flattened and fed into the five-layer FCNN that is same with the STFT case.

Each model uses the same train epoch of 600 samples (except for the SVM model). The evaluation results of each

TABLE I
Five different model architectures used for evaluation

Feature Extraction Methods		Classifier
STFT & 2D-CNN 1D-CNN Entropy		FCNN FCNN SVM
STFT & 2D-CNN & FCNN 1D-CNN & FCNN Entropy & FCNN	(Fusion) Linear / Tensor	FCNN

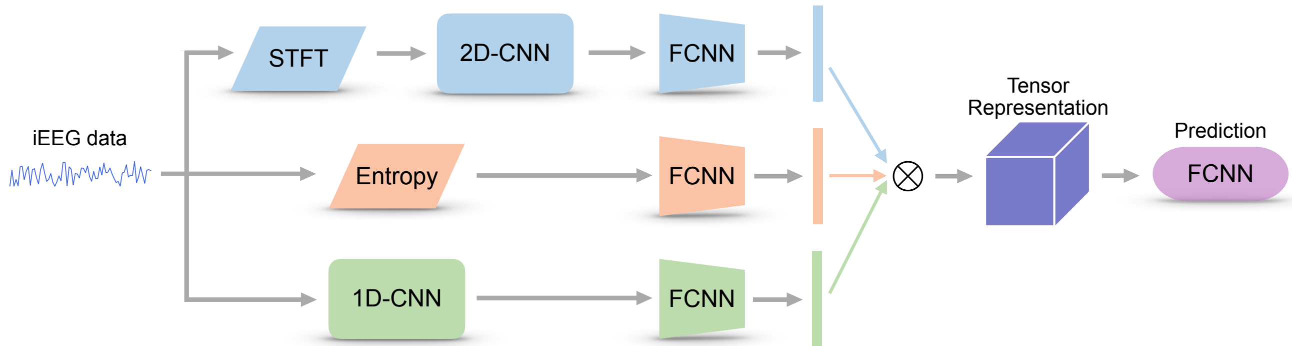

Fig. 4. The overall architecture of the proposed framework. The input iEEG signal is fed into all three sub-networks. The outputs of the three sub-networks are combined through tensor representation and fed into a fully connected network for prediction.

TABLE II
CLASSIFICATION RESULTS OF FOCUS AND NON-FOCUS IEEG DATA, COMPARING SINGLE FEATURE AND MULTI-FEATURE FUSION MODELS (MEAN $\pm$ STANDARD DEVIATION IN [%]).

Models	Accuracy	Precision	Recall	Specificity	F1-Score
Entropy	87.40 ± 0.6498	88.49 ± 1.070	85.98 ± 0.8412	88.81 ± 1.056	87.21 ± 0.7233
STFT	89.69 ± 0.03978	89.20 ± 0.1410	90.34 ± 0.09234	89.03 ± 0.1221	89.76 ± 0.04491
1D-CNN	89.80 ± 0.04611	92.29 ± 0.1670	86.83 ± 0.1078	92.75 ± 0.1223	89.47 ± 0.05684
Fusion (Linear)	90.84 ± 0.05912	94.92 ± 0.1646	86.34 ± 0.1116	95.33 ± 0.1260	90.39 ± 0.06753
Fusion (**Tensor**)	$\mathbf{93.44} \pm 0.03942$	94.28 ± 0.1110	92.50 ± 0.1260	94.38 ± 0.1371	93.38 ± 0.03876

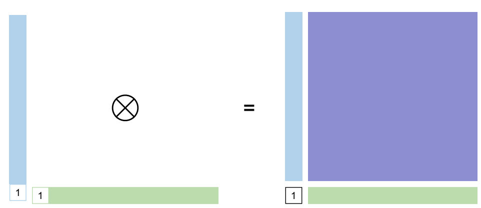

Fig. 5. Feature vector fusion with an additional '1' (two-dimensional graphical example).

model are shown in Table II, in which the mean and standard deviation of the last ten epochs are provided.

In the experiment results, multi-feature fusion model with tensor fusion shows the best performance. This could be explained by several reasons, among which the two main ones are: (a) compared to the single feature model, the fusion model considered three different features at the same time, while three features are interrelated and interacted with each other. In this way, the samples that cannot be classified well by the original feature sets individually may be possibly classified correctly by more expressive fused features. (b) compared to the linear fusion model, the tensor representation preserves the original multiple sets of features and also captures multilinear interactions between different sets of features, which can achieve better performance.

IV. Conclusion

In view of the problem of the long time and personal experience needed in the current diagnosis of epilepsy, and with the aim of improving the diagnostic process, an assistance system has been proposed that can reduce the workload of clinical experts. Furthermore, compared to the traditional method which only uses single type of features as the basis for judgement, the proposed method considers three types of features extracted from different angles. This process is similar to the diagnostic process of the clinical expert, who can analyze the iEEG signal comprehensively. According to experimental results, fusion models improve the performance compared to single-feature models. In addition, unlike the linear fusion model, cross-feature interactions between three types of features are performed using the tensor representation. With this approach, more effective features can be captured from iEEG signals and annotation information, which can significantly improve the classification performance as demonstrated by our extensive experiments.

Acknowledgments

This work was supported by JST CREST Grant Number JPMJCR1784 including AIP challenge program, Japan and JSPS KAKENHI (Grant No. 18K04178 and 20H04249). J.S-C. work is also based upon work from COST Action CA18106, supported by COST (European Cooperation in Science and Technology).

References

[1] T. Itakura and T. Tanaka, "Epileptic focus localization based on bivariate empirical mode decomposition and entropy," in *Proceedings of the Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, 2017, pp. 1426–1429.

[2] A. B. Das and M. I. H. Bhuiyan, "Discrimination and classification of focal and non-focal EEG signals using entropy-based features in the EMD-DWT domain," *Biomedical Signal Processing and Control*, vol. 29, pp. 11–21, 2016.

[3] O. Faust, U. R. Acharya, H. Adeli, and A. Adeli, "Wavelet-based EEG processing for computer-aided seizure detection and epilepsy diagnosis," *Seizure*, vol. 26, pp. 56–64, 2015.

[4] M. Li, W. Chen, and T. Zhang, "Classification of epilepsy EEG signals using DWT-based envelope analysis and neural network ensemble," *Biomedical Signal Processing and Control*, vol. 31, pp. 357–365, 2017.

[5] D. Chen, S. Wan, J. Xiang, and F. S. Bao, "A high-performance seizure detection algorithm based on discrete wavelet transform (DWT) and EEG," *PloS one*, vol. 12, no. 3, p. e0173138, 2017.

[6] R. Sharma, R. B. Pachori, and U. R. Acharya, "Application of entropy measures on intrinsic mode functions for the automated identification of focal electroencephalogram signals," *Entropy*, vol. 17, no. 2, pp. 669–691, 2015.

[7] A. Harati, M. Golmohammadi, S. Lopez, I. Obeid, and J. Picone, "Improved EEG event classification using differential energy," in *Proceedings of the IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, 2015, pp. 1–4.

[8] A. Bhattacharyya, R. B. Pachori, A. Upadhyay, and U. R. Acharya, "Tunable-Q wavelet transform based multiscale entropy measure for automated classification of epileptic EEG signals," *Applied Sciences*, vol. 7, no. 4, p. 385, 2017.

[9] N. E. Cámpora, C. J. Mininni, S. Kochen, and S. E. Lew, "Seizure localization using pre ictal phase-amplitude coupling in intracranial electroencephalography," *Scientific Reports*, vol. 9, no. 1, pp. 1–8, 2019.

[10] K. Edakawa, T. Yanagisawa, H. Kishima, R. Fukuma, S. Oshino, H. M. Khoo, M. Kobayashi, M. Tanaka, and T. Yoshimine, "Detection of epileptic seizures using phase–amplitude coupling in intracranial electroencephalography," *Scientific Reports*, vol. 6, p. 25422, 2016.

[11] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.

[12] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[13] N. Arunkumar, K. Ramkumar, V. Venkatraman, E. Abdulhay, S. L. Fernandes, S. Kadry, and S. Segal, "Classification of focal and non focal EEG using entropies," *Pattern Recognition Letters*, vol. 94, pp. 112–117, 2017.

[14] Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. H. Falk, and J. Faubert, "Deep learning-based electroencephalography analysis: a systematic review," *Journal of neural engineering*, vol. 16, no. 5, p. 051001, 2019.

[15] R. G. Andrzejak, K. Schindler, and C. Rummel, "Nonrandomness, nonlinear dependence, and nonstationarity of electroencephalographic recordings from epilepsy patients," *Physical Review E*, vol. 86, no. 4, p. 046206, 2012.

[16] A. Medvedev, G. Agoureeva, and A. Murro, "A long short-term memory neural network for the detection of epileptiform spikes and high frequency oscillations," *Scientific Reports*, vol. 9, no. 1, pp. 1–10, 2019.

[17] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces," *Journal of neural engineering*, vol. 15, no. 5, p. 056013, 2018.

[18] A. Zadeh, M. Chen, S. Poria, E. Cambria, and L.-P. Morency, "Tensor fusion network for multimodal sentiment analysis," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2017, pp. 1103–1114.