Power of Deep Learning for Channel Estimation Signal Detection in OFDM Systems and Rehabilitation of Traumatic Brain Injured Patients

Veena S[1], Suraj kumar[2], Sankalp Kumar Rai[3]
Department of Computer Science Engineering
[1] Assistant Professor, Atria Institute of Technology, Visvesvaraya Technological University, Bangalore
[2,3] Student, Atria Institute of Technology, Visvesvaraya Technological University, Bangalore

Abstract- *This letter presents our initial results in deep learning for channel estimation and signal detection in orthogonal frequency-division multiplexing (OFDM) systems and Rehabilitation of Traumatic Brain Injured Patients. Different from existing OFDM receivers that first estimate channel state information (CSI) explicitly and then detect/recover the transmitted symbols using the estimated CSI, the proposed deep learning-based approach estimates CSI implicitly and recovers the transmitted symbols directly. To address channel distortion, a deep learning model is first trained offline using the data generated from simulation based on channel statistics and then used for recovering the online transmitted data directly.*
we have analyzed the methods for facial features extraction for TBI patients to determine optimal time to have aforementioned rehabilitation process on the basis of positive and negative facial expressions. We have employed a deep learning architecture based on convolutional neural network and long short term memory on RGB and thermal data that were collected in challenging scenarios from real patients. It automatically identifies the patient's facial expressions, and inform experts or trainers that "it is the time" to start rehabilitation session.

Keywords- Deep learning ,system Architecture and CNN + LSTM Architecture , METHDOLOGY, RELATED WORK, CONCLUSION.

I. INTRODUCTION

ORTHOGONAL frequency-division multiplexing (OFDM) is a popular modulation scheme that has been widely adopted in wireless broadband systems to combat frequency-selective fading in wireless channels. Channel state information (CSI) is vital to coherent detection and decoding in OFDM systems. Usually, the CSI can be estimated by means of pilots prior to the detection of the transmitted data. With the estimated CSI, transmitted symbols can be recovered at the receiver. Historically, channel estimation in OFDM systems has been thoroughly studied. The traditional estimation methods, i.e., least squares (LS) and minimum mean-square error (MMSE),have been utilized and optimized in various conditions [2]. The method of LS estimation requires no prior channel statistics but its performance may be inadequate. The MMSE estimation in general leads to much better detection performance by utilizing the second order statistics of channels.

Traumatic brain injury (TBI) causes life-long damage to cognitive, physical, behavioural and social functions. It may take up to 5 years or more for recovery after TBI [1]. According to International Brain Injury Association (IBIA), annually one million people suffer from traumatic brain injury (TBI) only in America whereas same number of people suffer with TBI in Europe [2]. American Center for Disease Control and Prevention estimates more than 3.7 million people are living with long term disability after TBI [3]. During rehabilitation period, patient has to live in a specialized care center called neuro-center or care home where the main focus is on the retraining of activities of daily life, cognitive, social and physical exercises through a set of protocols. Recovery targets are based on determination of combination of cognitive, behavioral and physical shortfalls. It is seen that rehabilitation activities are performed daily on set time table of neuro-center, regardless of mental conditions of subject. This leads to more time expensive training with less result oriented outcome.

II. DEEP LEARNING BASED ESTIMATION AND DETECTION

B.a Deep Learning Methods

Deep learning has been successfully applied in a wide range of areas with significant performance improvement, including computer vision [6], natural language processing[7], speech recognition [8], and so on. A comprehensive introduction to deep learning and machine learning can be found in [1].The structure of a DNN model is shown in Fig. 1. Generally speaking, DNNs are deeper versions of ANNs by increasing the number of hidden layers in order to improve the ability.

Flow diagram

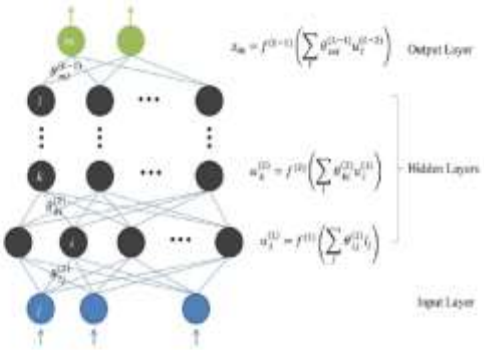


Fig. 1. An example of deep learning models.

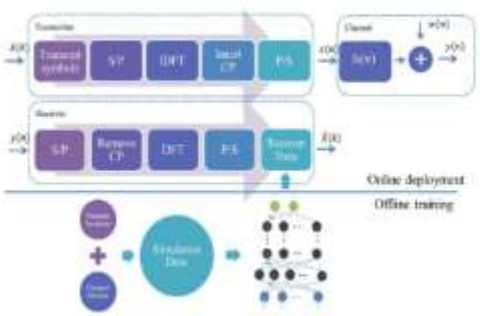


Fig. 2. System model.

in representation or recognition. Each layer of the network consists of multiple neurons, each of which has an output that is a nonlinear function of a weighted sum of neurons of its preceding layer, as shown in Fig. 1. The nonlinear function may be the Sigmoid function, or the Relu function, defined as $f_S(a) = 1/(1+e^{-a})$, and $f_R(a) = \max(0, a)$, respectively. Hence, the output of the network $\mathbf{z}$ is a cascade of nonlinear transformation of input data $\mathbf{I}$, mathematically expressed as

$$\mathbf{z} = f(\mathbf{I}, \boldsymbol{\theta}) = f^{(L-1)}(f^{(L-2)}(\cdot \cdot \cdot f^{(1)}(\mathbf{I}))), (1)$$

where L stands for the number of layers and $\boldsymbol{\theta}$ denotes the weights of the neural network.

B.b System Architecture

The architecture of the OFDM system with deep learning based channel estimation and signal detection is illustrated in Fig. 2. The baseband OFDM system is the same as the conventional ones. On the transmitter side, the transmitted symbols inserted with pilots are first converted to a paralleled data stream, then the inverse discrete Fourier transform (IDFT) is used to convert the signal from the frequency domain to the time domain. After that, a cyclic prefix (CP) is inserted to mitigate the inter-symbol interference (ISI). The length of the CP should be no shorter than the maximum delay spread of the channel.We consider a sample-spaced multi-path channel described by complex random variables $\{h(n)\}^{N-1}_{n=0}$. The received signal can be expressed as

$$y(n) = x(n) \otimes h(n) + w(n),$$

CNN + LSTM Architecture

After the pre-processing of the data, it is fed to 2D-CNN for training purpose for mood recognition based on PE and NE. This network is fine tuned by VGG-16 face model [35] for spatial feature extraction. CNN parameters are initialized randomly and through back propagation using gradient descent its weights are adjusted. Thermal data is also fine tuned with pre-trained VGG-16 face (RGB) model. CNN deals with frames in isolated manner. For capitalizing on relation with time, special Recurrent Neural Network (RNN) called LSTM is employed. LSTM is gate controlled network with input (i), output(o) and forget (f) gates. LSTM gates holds the input information as long as its forget gate is not triggered to acquire the temporal information between frames for said purposes. These gates control the flow of instructions by point wise multiplication and sigmoid functions σ, which bound the information flow between zero and one by the followings:

$$i(t) = \sigma(W(x \to i)x(t) + W(h \to i)h(t-1) + b(1 \to i)) \quad (1)$$

$$f(t) = \sigma(W(x \to f)x(t) + W(h \to f)h(t-1) + b(1 \to f)) \quad (2)$$

In these equations, W are weights associated with activated neurons for particular input i. Where as σ squashes the value of activation between the range of 0 and 1.

Fig. 1. CNN-LSTM based deep learning architecture for both modalities to exploit spatio-temporal information for FER.

III. METHDOLOGY

This section presents the architecture of the intended approach for FER analysis of real TBI patient in realistic environment. We have employed the same method as followed in [5] but employed new pre-processing technique of face frontalization because of large pose variation. We tested the deep learning method of [5] on both modalities, with early and late fusions. Facial expressions are recognized by employing CNN (to use spatial features) and linking with LSTM to utilize spatio-temporal attributes of RGB, thermal and fused RGB thermal modalities. The block diagram of the proposed method is illustrated in Figure 1. The steps of the proposed system are further explained in the following subsections.

Block diagram

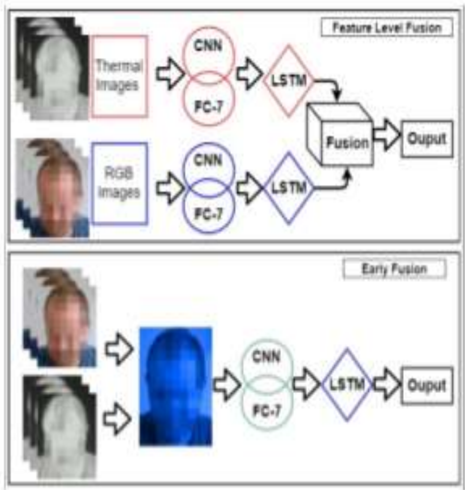


Fig. 2. Block diagram of early and Feature Level Fusion of modalities for FER.

In order to analyze the ability of both modalities in FER applications, two approaches were employed: 1) data level fusion (early) 2) feature level Fusion. In the first approach both modalities are combined into data array for feature learning through CNN. In the second method, both RGB and thermal imagery features are fed separately into deep learning system for feature learning and combined together as input for second classifier (LSTM) for final output. Block diagram of both modalities can be seen in Figure 2.

IV.RELATED WORK

Current FER system can be categorized on the basis of methods used for feature extraction and classification. Our main focus is on the methods involving Convolution Neural Networks (CNN) or other deep learning approaches as they provide state of the art results for, e.g., face recognition [14] [15] [16], facial expressions recognition [17] [18] [19] [20] [21] [22] [23] [12] [13] and emotional states identification [24] [25] [26] [27]. Handcrafted features such as Local Binary Pattern (LBP), SIFT, Local Quantized Pattern (LPQ) and Histogram of Oriented Gradients (HOG) applied in [28] [29][30][31][32] are outperformed by CNN based deep neural networks despite their low computational cost. In [17], Tang proposed deep CNN along with Support Vector Machines (SVM) and achieved state of the arts results for FERwith1stprizeinFER-2013competition.In2014,Liu[19] performed three functions- feature learning, feature selection and classification in unified manner through Boosted Deep Belief Networks (BDBN). This method worked exceptionally well even for extremely complicated features from facial image. [22] used DBN models to overcome the limitations of linear feature selections. Yu and Zang [20] in 2015, presented their work for Emotion recognition in Wild challenge for image based static FER. They have applied multiple deep CNN with random initialization of each network and minimized likelihood and hinge loss. Their results surpassed the challenge baseline significantly. In year 2017, [13] exercised CNN to learn features from VGG-Faces and integrated with Long Short Term Memory (LSTM) to gain the temporal information. This approach was further improved by[12] who applied deep CNN for features classification into expressions and feed the system with super-resolved facial images.

E. OFDM Model Training

The models are trained by viewing OFDM modulation and the wireless channels as black boxes. Historically, researchers have developed many channel models that well describe the real channels in terms of channel statistics. With these channel models, the training data can be obtained by simulation. In each simulation, a random data sequence is first generated as the transmitted symbols and the corresponding OFDM frame is formed with a sequence of pilot symbols and the pilot symbols need to be fixed during the training and deployment stages. The current random channel is simulated based on the channel models. The received OFDM signal is obtained based on the OFDM frames undergoing the current channel distortion, including the channel noise. The received signal and the original transmitted data are collected as the training data. The input of deep learning model is the received data of the pilot block and one data block. The model is trained to minimize the difference between the output of the neural network and the transmitted data. The difference can be portrayed in several ways.

In our experiment settings, we choose the *L*2 loss,

$$L_2 = \frac{1}{N}\sum_{k}(\hat{X}(k) - X(k))^2,$$

where ˆ*X (k)* is the prediction and *X(k)* is the supervision message, which is the transmitted symbols in this situation. The DNN model we use consists of five layers, three of which are hidden layers. The numbers of neurons in each layers are 256, 500, 250, 120, 16, respectively. The input number corresponds to the number of real parts and imaginary parts of 2 OFDM blocks that contain the pilots and transmitted symbols. Every 16 bits of the transmitted data are grouped and predicted based on a single model trained independently, which is then concatenated for the final output. The Relu function is used as the activation function in most layers except in the last layer where the Sigmoid function is applied to map the output to the interval [0, 1].

TBI PATIENT DATABASE FOR FER

To analyze facial expressions, data is collected in three prespecified scenarios from seven TBI patients in two modalities: RGB and Thermal. Pre-specified scenarios in data collection are maintained to have reliable data for further use. Those scenarios are:

1) **cognitive activity**
2) **physiotherapy and**
3) **social communication.**

These scenarios are selected after consulting many experts and care givers, who are working on rehabilitation of TBI individuals in Denmark. On contrary to healthy people, as mentioned in [5], data acquisition task is quite complicated due to extreme behavioural responses, verbalization, physical aggression, impaired reasoning, reduced cognitive skills along with frequent pose variations. Ilyas et al. [5], collected RGB database by Axis RGB-Q16 camera with resolution of 1280 x 960 to 160 x 90 pixels at 30fps (frames per second) and applied pre processing techniques of face detection, FQA, (Supervised Decent Method) SDM for landmark detection and tracking before logging into a face log. We have operated with a Logitech camera as well to record the starting and ending time stamp of particular expressions. Along with RGB, we have gathered thermal images of TBIsubjectswithAxisThermal-Q1922camerawith focal lens of 10 mm.

TABLE I
DATABASE OF TBI PATIENTS WITH ACTIVITY PARTICIPATION

Subjects	Number of Sessions	Activities Participated		
		Cognitive	Social Comm	Physiotherapy
Subject A	7	Y	Y	Y
Subject B	5	Y	Y	Y
Subject C	5	Y	Y	Y
Subject D	7	Y	X	Y
Subject E	3	Y	X	X
Subject F	4	Y	Y	Y
Subject G	3	X	Y	Y

APPLICATION:

A.Impact of Pilot Numbers
B. Impact of CP
C. Impact of Clipping and Filtering Distortion
D. Robustness Analysis

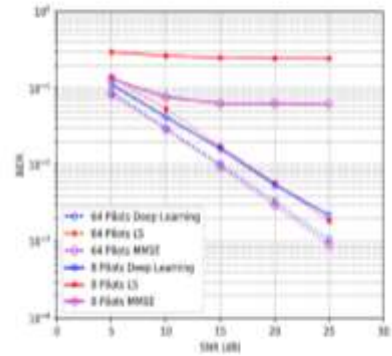
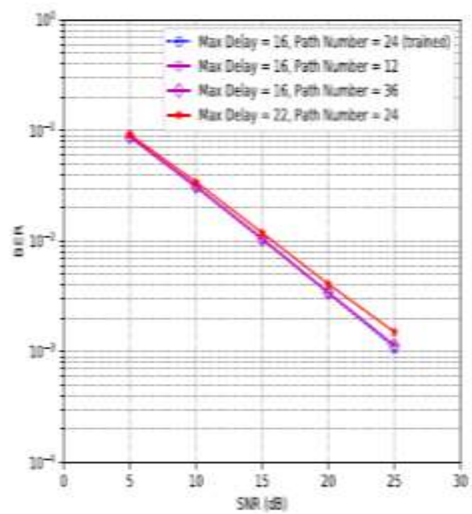


Fig. 3. BER curves of deep learning based approach and traditional methods.

In the simulation above, the channels in the online deployment stage are generated with the same statistics that are used in the offline training stage. However, in real-world applications, mismatches .may occur between the two stages. Therefore, it is essential for the trained models to be relatively robust to these mismatches. In this simulation, the impact of variation in statistics of channel models used during training and deployment stages is analyzed. Fig. 7 shows the BER curves when the maximum delay and the number of paths in the test stage vary from the parameters used in the training stage described in the beginning of this section. From the figure, variations on statistics of channel models do not have significant damage on the performance of symbol detection.



BER curves with mismatches between training and deployment stages.

V. EXPERIMENTAL RESULTS

We demonstrate the results in the following contexts: a) Classification of six basic expression groups in both early and feature level fusion scenarios to evaluate the performance of CNN+LSTM based FER

b) PE and NE classifications before and after face frontalization on all individual modalities and fusions.

TABLE II
RECOGNITION ACCURACY OF PROPOSED METHOD IN DIFFERENT CONTEXTS

Confusion Matrix %	RGB Non Frontal		RGB Frontal		Thermal		Early Fusion		[13]		Feature Level Fusion	
	PE	NE	PE	NE	PE	NE	PE	NE	PE	NE	PE	NE
Positive Expression (PE)	0.75	0.17	0.86	0.15	0.69	0.25	0.84	0.14	0.79	0.12	0.86	0.11
Negative Expression (NE)	0.25	0.71	0.11	0.87	0.23	0.65	0.16	0.79	0.1	0.82	0.09	0.89
Recognition Accuracy (%)	79.34		86.85		74.65		84.39		87.97		88.74	

TABLE IV
MCONFUSION MATRIX BY FEATURE LEVEL FUSION OF MODALITIES FOR 6 BASIC FER.

	Neutral	Happy	Angry	Sad	Fatigued	Surprised
Neutral	0.77	0.03	0.02	0.07	0.07	0.01
Happy	0.04	0.71	0.02	0.03	0.05	0.16
Angry	0.04	0.02	0.81	0.09	0.03	0.02
Sad	0.07	0.01	0.05	0.76	0.13	0.01
Fatigued	0.09	0.01	0.09	0.1	0.55	0.11
Surprised	0.07	0.14	0.1	0.02	0.06	0.56

FUTURE SCOPE

- We can implement various bioelectric sensors on the helmet to measure various activity.
- We can use small camera for the recording the drivers activity.
- It can be used for passing message from the one vehicle to another vehicle by using wireless transmitter.
- We have used solar panel for helmet power supply by using same power supply we can charge our mobile.

VI. CONCLUSION

we have demonstrated our initial efforts to employ DNNs for channel estimation and symbol detection in an OFDM system. The model is trained offline based on the simulated data that view OFDM and the wireless channels as black boxes. The simulation results show that the deep learning method has advantages when wireless channels are complicated by serious distortion and interference, which proves that DNNs have the ability to remember and analyze the complicated characteristics of the wireless channels. For real-world applications and Mood recognition is important task for rehabilitation and care centers. In this work we have faced the challenge of mood recognition of TBI patients rather than facial expression recognition for healthy people. In case of TBI individuals, extraction of all expression is very complicated and its dependant to patient disability and FER did not provide good results [5]. However, we recognized the mood of patients with accuracy of 86.93 percentage that is very close to [13] system when implemented on TBI patient database. So this system can help physiotherapist and trainers in fast rehabilitation process after recognizing the positive mood of the patient. Furthermore, we applied early and feature level fusion to enhance the recognition rate of the system. Our system results can be improved further by employing 3D face frontalization. Even though the results are encouraging, efforts are still in progress to provide the robust solutions to deal with real time and environment challenges like real time computation or patient positioning.

REFERENCES

[1] J. Schmidhuber, “Deep learning in neural networks: An overview,”Neural Netw., vol. 61, pp. 85–117, Jan. 2015.

[2] Y. G. Li, L. J. Cimini, and N. R. Sollenberger, “Robust channel estimationfor OFDM systems with rapid dispersive fading channels,” IEEETrans. Commun., vol. 46, no. 7, pp. 902–915, Jul. 1998.

[3] X. Wang, L. Gao, S. Mao, and S. Pandey, “CSI-based fingerprintingfor indoor localization: A deep learning approach,” IEEE Trans. Veh.Technol., vol. 66, no. 1, pp. 763–776, Jan. 2017.

[4] Brain injury facts—international brain injury association-ibia. [Online]. Available: http://www.internationalbrain.org/brain-injury-facts/

[5] J. Bird and R. Parente, Recognition of nonverbal communication of emotion after traumatic brain injury, 2014..

[6] B. Dang, W. Chen, W. He, and G. Chen, “Rehabilitation treatment and progress of traumatic brain injury dysfunction,” Neural Plasticity, vol. 2017, 2017.

[7] E. Nachmani, Y. Be’ery, and D. Burshtein, “Learning to decodelinear codes using deep learning,” in Proc. 54th Annu. AllertonConf. Commun. Control Comput., Monticello, IL, USA, Sep. 2016,pp. 341–346.

[8] M. P. Hyett, G. B. Parker, and A. Dhall, The Utility of Facial Analysis Algorithms in Detecting Melancholia. Cham: Springer International Publishing, 2016, pp. 359–375.

[9] M. Bellantonio, M. A. Haque, P. Rodriguez, K. Nasrollahi, T. Telve, S. Escalera, J. Gonzalez, T. B. Moeslund, P. Rasti, and G. Anbarjafari, Spatio-temporal Pain Recognition in CNN-Based Super-Resolved Facial

Images. Cham: Springer International Publishing, 2017, pp. 151–162.