Disentangled Adversarial Autoencoder for Subject-Invariant Physiological Feature Extraction

Mo Han[1], Ozan Özdenizci[1], Ye Wang[2], Toshiaki Koike-Akino[2] and Deniz Erdoğmuş[1]

***Abstract*—Recent developments in biosignal processing have enabled users to exploit their physiological status for manipulating devices in a reliable and safe manner. One major challenge of physiological sensing lies in the variability of biosignals across different users and tasks. To address this issue, we propose an adversarial feature extractor for transfer learning to exploit disentangled universal representations. We consider the trade-off between task-relevant features and user-discriminative information by introducing additional adversary and nuisance networks in order to manipulate the latent representations such that the learned feature extractor is applicable to unknown users and various tasks. Results on cross-subject transfer evaluations exhibit the benefits of the proposed framework, with up to 8.8% improvement in average accuracy of classification, and demonstrate adaptability to a broader range of subjects.**

***Index Terms*—adversarial deep learning, stress assessment.**

I. Introduction

RECENTLY, biosignal processing has obtained increasing significance, since the abilities of machines to understand human emotions, discern physiological disorders, and execute appropriate actions are key points in the area of human computer interaction (HCI) [1]. HCI enables users to communicate their physiological information with machines for help with manipulating external devices in a more reliable, robust and safe manner. Traditionally, the assessment of physiological activity (e.g., human stress level and mental status) was implemented by monitoring signals such as electroencephalography (EEG) [2] and electromyography (EMG) [3]. However, these measurements require either surface (non-invasive) or implanted (invasive) electrodes and frequent calibration, which increase system cost and decrease user comfort. To realize more portable interfaces, significant progress was recently achieved with wearable sensors for precisely monitoring physiological signals such as heart rate, skin temperature, and arterial oxygen level [4–9]. These more convenient (non-EEG) biosignals avoid the aforementioned issues, and can be obtained from a wrist-worn platform in more effective, comfortable, and less expensive ways.

However, biosignals often vary across subjects and recording sessions of the same person depending on physical/mental conditions or the disturbance by task-irrelevant activity. Such variability is an obstacle to successful HCI applications to a wider range of users and tasks, since biosignals are often collected from limited number of subjects. Under such restrictions, a robust feature extractor can be constructed with transfer learning [10–12], which tries to discover shared data features that are invariant across subjects and tasks. In particular, promising results were demonstrated for transfer learning by censoring nuisance features via adversarial training [13–19]. These works use adversarial methods to learn universal features shared by an attribute group, where a discriminative unit distinguishes shared features with respect to the different attributes adversarially to the feature extractor. However, in existing works, the adversarial unit will act directly on the entire latent representation to preserve cross-attribute shared features, leading to loss of attribute-specific information. Hence, instead of simply reserving shared features only with one adversarial discriminator, we disentangle the physiological latent representations into 2 parts of subject- and task-relevant features by jointly training two discriminators, so that the model can better handle both subject- and task-specific variations.

This paper proposes an extended adversarial feature encoding to exploit disentangled universal representations, motivated by [17], where the adversarial classifier is generalized into a feature extractor. Unlike classic feature extractors ignoring the target subjects and task calibration, we introduce two additional networks, i.e., adversary and nuisance blocks, in an autoencoder (AE) architecture to re-organize the latent representations, thereby accounting for a trade-off between task-related features and person-discriminative information. Even if a new user is dissimilar to any of the training subjects, instead of reconciling to suboptimality by extracting subject-specific features only, task-relevant representations can still be incorporated into the feature extraction. Empirical assessments were performed on a publicly available dataset of physiological biosignals for human stress level assessment. Results show a significant advantage of the disentangled adversarial framework through cross-subject evaluations with various classifiers, achieving up to 8.8% improvement in classification accuracy.

II. Methods

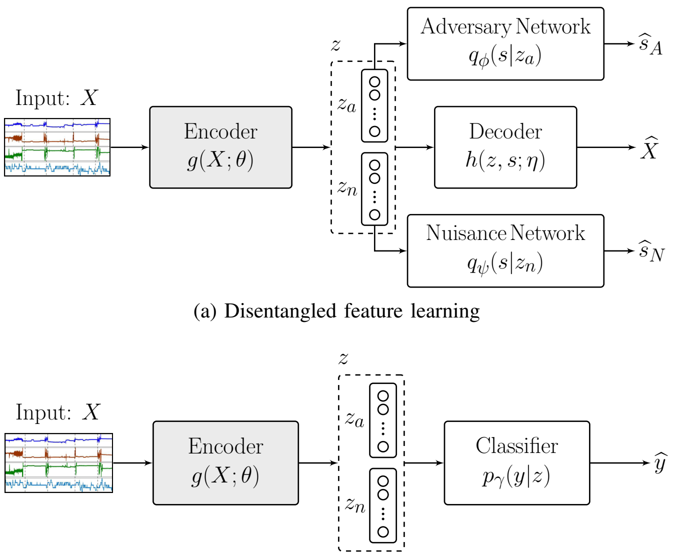
A. Disentangled Adversarial Transfer Learning

Define $\{(X_i, y_i, s_i)\}_{i=1}^{n}$ as a training set, where $X_i \in \mathbb{R}^C$ is the data matrix of trial i recorded from C channels, $y_i \in \{0, 1, \dots, L-1\}$ is the label of physiological status/task among L classes, and $s_i \in \{1, 2, \dots, S\}$ is the identification (ID) number among S subjects. We assume the label y and subject ID s are marginally independent, and the data is

[1] Cognitive Systems Laboratory, Department of Electrical and Computer Engineering, Northeastern University, Boston, MA 02115, USA, E-mail: {han, oozdenizci, erdogmus}@ece.neu.edu.

[2] Mitsubishi Electric Research Laboratories (MERL), Cambridge, MA 02139, USA, E-mail: {yewang, koike}@merl.com.

M. Han was an intern at MERL during this work. O. Özdenizci and D. Erdoğmuş are partially supported by NSF (IIS-1149570, CNS-1544895, IIS-1715858), DHHS (90RE5017-02-01), and NIH (R01DC009834).



(a) Disentangled feature learning

(b) Task-specific classifier training with frozen encoder

Fig. 1: (a) An encoder $z = g(X;\theta)$ and decoder $\hat{X} = h(z,s;\eta)$ conditioned on s is trained to learn a latent variable z from data X. Latent z is divided to z_a and z_n, where z_a is input to an *adversary network*, while z_n is fed to a *nuisance network*. The full latent vector $z = [z_a, z_n]$ is used as an input to the decoder, alongside the condition s. (b) Using the pre-trained encoder frozen as a static feature extractor to generate z form X, a classifier is then optimized to predict the corresponding user physiological status or any other task.

dependent on y and s, i.e., $X \sim p(X|y,s)$. Our goal is to build a model that estimates the label y of a given observation X, where the model is robust to the variability across subjects s, which captures the nuisance variations we wish to suppress for transferring feature extraction. In the proposed method, we first train a disentangled feature extractor based on a modified AE (i.e., an encoder-decoder pair), and then utilize this learned encoder as a static feature extractor to further train a task classifier for the final discriminative model as shown in Fig. 1.

AEs are feature learning machines which constitute an encoder and decoder network pair. The encoder learns a latent vector to represent data features, while the decoder aims to recover the input data from this learned latent representation. Here, we train a modified AE to extract the latent feature z from data X via the encoder $z = g(X;\theta)$ with parameters θ, attached with a decoder $\hat{X} = h(z,s;\eta)$ parameterized by η, where the decoder is conditioned on the nuisance variable s as an additional input along with z. The conditional decoder output $\hat{X} \in \mathbb{R}^C$ is a reconstructed estimate of input X.

In the proposed model, the latent representation z consists of two sub-parts: z_a and z_n, divided by the ratio of $(1-r_N) : r_N$ over their dimensionality. The representation z_a is fed into an *adversary network* with parameters ϕ, while the feature z_n is input to another *nuisance network* parameterized by ψ, as illustrated in Fig. 1(a). The full latent space $z = [z_a, z_n]$ is further fed into the decoder $h(z,s;\eta)$, which is conditioned on s. Disentangling z into sub-parts z_a and z_n is proposed to systematically re-arrange the features related to task and subject respectively: z_a aims to conceal the subject information regarding s, while z_n is designed to include the subject-related features. By dissociating the nuisance variable from task-related features, the model is extrapolated into a broader domain of subjects. For an unknown user, the subject-invariant feature z_a would be useful for the final prediction; simultaneously, the biosignal which is similar to known subjects can also be projected to z_n as an additional side information. In order to embed more task-related information into z_a and filter out factors of variation caused by s, the encoder is forced to minimize the likelihood $q_\phi(s|z_a)$ of z_a; on the other hand, to retain sufficient subject-related information within z_n, the encoder is simultaneously designed to maximize the likelihood $q_\psi(s|z_n)$ from z_n. The encoder-decoder pair is trained to minimize the reconstruction loss between X and $\hat{X}$. Hence, the overall loss to train the proposed model is given by

$$\begin{aligned}\mathsf{Loss}_{\mathrm{AE}}(X;\eta,\theta,\psi,\phi) &= -\mathbb{E}\big[\log p_\eta\big(\hat{X}|g(X;\theta),s\big)\big] \\ &\quad - \lambda_N \mathbb{E}\left[\log q_\psi(s|z_n)\right] + \lambda_A \mathbb{E}\left[\log q_\phi(s|z_a)\right], \qquad (1)\end{aligned}$$

where the first term is the reconstruction loss of the decoder $\hat{X} = h(z,s;\eta)$ with $z = g(X;\theta)$, and λ_A and λ_N denote the weight parameters for adversary and nuisance networks respectively, to implement an adjustable trade-off between invariance and identification performance. When $\lambda_A = \lambda_N = 0$, the model reduces to a regular conditional AE (cAE) structure without the disentangling transfer learning units.

In addition to the overall objective, at each training iteration, the adversary and nuisance networks are optimized towards predicting the variable s among S subjects by maximizing the likelihoods $q_\phi(s|z_a)$ and $q_\psi(s|z_n)$ respectively. For the parameter updates at each iteration, optimization is performed by stochastic gradient descent alternatingly among the adversary network, nuisance network and encoder-decoder pair, where the adversary and nuisance networks are individually trained to minimize their cross-entropy losses.

Attached to the pre-trained disentangled encoder whose network weights are frozen, a separate classifier is then trained using the feature representation z, as shown in Fig. 1(b). The task classifier aims to predict the user physiological status or task category y given observation X among L classes, where X would pass through the feature extractor $z = g(X;\theta)$ before the task classifier. Classifier training is performed to minimize the softmax cross-entropy loss: $\mathbb{E}\left[-\log p_\gamma(\hat{y}|z)\right]$, where $\hat{y}$ is the task category estimate, and γ are task classifier parameters.

B. Model Architecture

Deep learning frameworks have shown promising performance in biosignal processing recently [15, 20, 21]. In the light of these works, we mainly focus on neural network feature extractor. We however note that any other discriminative learning methods can be used in the proposed methodology of disentangled adversarial transfer learning. The model architecture specifications we used in our experiments are presented in Table I. Latent representation z with dimensionality $d = 15$ is generated and split into sub-representations z_n and z_a with dimensions of $d \cdot r_N$ and $d \cdot (1 - r_N)$, which are respectively

TABLE I: Network architectures; FC($d_\mathrm{i}, d_\mathrm{o}$): fully connected linear layer with input/output dimensions d_i and d_o, ReLU: rectified linear unit

Encoder Network	FC(C, 15) → ReLU → FC(15, 15)
Decoder Network	FC(15, 15) → ReLU → FC(15, C)
Adversary Network	FC(15, S)
Nuisance Network	FC(15, S)

fed into adversary and nuisance networks with the same output dimensionality S for the classification of subject IDs.

To verify the adaptability of the proposed feature extractor framework with adversarial disentangling, multiple structures for task classification were implemented, including multilayer perceptron (MLP), nearest neighbors, decision tree, linear discriminant analysis (LDA), linear support vector machine (SVM), and logistic regression classifiers.

III. Experimental Evaluation

A. Physiological Biosignal Dataset

We evaluate our model on a publicly available biosignal dataset for assessment of stress status levels [4]. This database consists of multi-modal physiological biosignals for inferring $L = 4$ discrete stress status levels from $S = 20$ healthy subjects, including physical stress, cognitive stress, emotional stress and relaxation. The data were collected by non-invasive wrist-worn biosensors, measuring electrodermal activity, temperature, three-dimensional acceleration, heart rate, and arterial oxygen level. Thus, the data consist of signals from $C = 7$ channels in total, which were temporally downsampled to 1 Hz sampling rate to align all data channels. Each subject performed 7 trials, where 4 out of the 7 trials were for the relaxation status, over a measurement session lasting approximately 35 minutes. To account for imbalanced number of trials across classes, we excluded the last three relaxation trials, resulting in one trial per stress status.

B. Experiment Implementation

The regularization parameters λ_A and λ_N, as well as the ratio of nuisance feature r_N were to be determined. The model was trained with different parameter combinations, and favored decreasing in adversary accuracy with increasing nuisance accuracy, while achieving high task classifier accuracy on validation sets. To narrow down the amount of λ_A and λ_N parameter combination options, we first set $\lambda_A = 0$ to optimize λ_N; then froze λ_N at its optimized value from the previous step to choose an optimal λ_A. The value ranges we used for these parameters are $\lambda_A \in \{0, 0.01, 0.1, 0.2, 0.5\}$ and $\lambda_N \in \{0, 0.005, 0.01, 0.2, 0.5\}$. Based on our model that the user-related features z_n will not vary dramatically across different users, we fixed the ratio of nuisance representation to $r_N = 1/3$. Note that these parameter combinations can be further optimized by cross-validating the model learning process. Evaluations were performed by cross-subjects transfer analysis using a leave-one-subject-out approach, where the left-out subject constituted the cross-subject test set, and the training and validation sets were composed of 90% and 10% randomized trial splits from the remaining subjects.

C. Results and Discussion

Transfer learning accuracies for 20 held-out subjects with different classifiers and feature learning models are presented in Fig. 2. Specifically, AE is a baseline encoder-decoder pair whose decoder is $h(z; \eta)$, cAE is a conditional AE whose decoder $h(z, s; \eta)$ is conditioned on s, A-cAE and D-cAE are cAE with only an adversary or nuisance network present respectively, and DA-cAE specifies our cAE with both adversary and nuisance networks. Corresponding parameter choices for each classifier with each model in Fig. 2 are presented in Table II, which were optimized via a parameter sweep as previously described. Note that the A-cAE corresponds to the adversarial learning methods presented in [14, 16, 18, 19].

As observed in Fig. 2 and Table II, when we compare cAE and AE models, simply providing the conditional input s to the decoder can slightly improve the classification performance. We further observe increased accuracies with A-cAE and D-cAE models compared to cAE, indicating that more subject-shared information remaining in z_a results in better decoding of y. More importantly, DA-cAE achieves a further improvement of up to 8.8% in an average accuracy compared to the regular AE, and also outperforms individual regularization approaches A-cAE and D-cAE. With both adversary and nuisance networks, our feature extractor leads to more stable performance universally across all subjects and all task classifiers. In addition, the worst-case transfer accuracies are highly improved as observed in Fig. 2, showing that the proposed transfer learning framework yields better robustness to novel users from a broader range by disentangling subject- and task-related representations at the feature extractor end.

We here focus on the MLP task classifier to discuss the impact of parameters in DA-cAE. As shown in Table III, we first assessed the baseline AE and cAE with $\lambda_A = \lambda_N = r_N = 0$ to train the MLP task classifier. Then, we evaluated the D-cAE with $\lambda_N \in \{0.005, 0.01, 0.2, 0.5\}$, $\lambda_A = 0$ and $r_N = 1/3$. Finally, we fixed $\lambda_N = 0.005$ to evaluate our DA-cAE with different choices of $\lambda_A \in \{0.01, 0.1, 0.2, 0.5\}$. For each of these parameter combinations, Table III shows the average accuracies of the MLP task classifier (i.e., 4-class stress level decoding), as well as the accuracy of adversary and nuisance networks (i.e., 20-class ID decoding). A higher accuracy of the main MLP task classifier indicates better discrimination of stress levels, a higher accuracy of nuisance network implies that more subject-dependent features are inherent in the representation z_n, and a lower accuracy of adversary network demonstrates that subject-invariant task-specific information are preserved in the learned representation z_a. We notice that with increasing λ_N, the nuisance network accuracy grows and specifically $\lambda_N = 0.005$ with $r_N = 1/3$ leads to higher task classification accuracy. Moreover, when fixing $\lambda_N = 0.005$ and $r_N = 1/3$, we observe that higher λ_A censors the encoder with decreased adversary network accuracy, and therefore enforces stronger task-specific information but less extraction of subject-relevant information into the learned z_a.

The convergence curve of the optimized DA-cAE case from Table III is shown in Fig 3, where DA-cAE loss (1) converges within 5 epochs. With more training epochs, the loss value

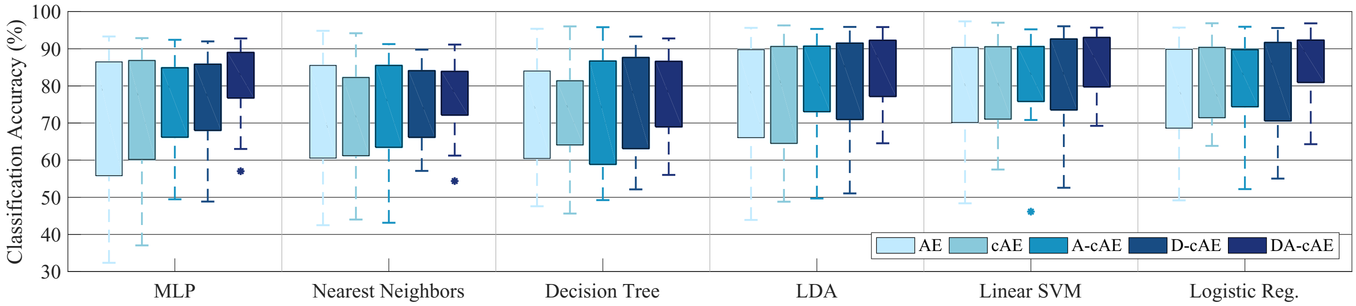


Fig. 2: Transfer learning accuracies for 20 held-out subjects: (1) regular AE with decoder $h(z;\eta)$, (2) cAE: AE with s-conditional decoder $h(z,s;\eta)$, (3) A-cAE: cAE with an extra adversary network, (4) D-cAE: cAE with an extra nuisance network, (5) DA-cAE: cAE with both adversary and nuisance networks. For each box, the central line marks the median, upper and lower bounds represent first and third quartiles, and dashed lines denote extreme values.

TABLE II: Optimized parameters and corresponding averaged cross-subject accuracies on five models for six classifiers

	AE				cAE				A-cAE [14, 16, 18, 19]				D-cAE				**DA-cAE (Proposed)**			
	λ_A	λ_N	r_N	avg acc	λ_A	λ_N	r_N	avg acc	λ_A	λ_N	r_N	avg acc	λ_A	λ_N	r_N	avg acc	λ_A	λ_N	r_N	avg acc
MLP	0	0	0	72.2%	0	0	0	72.9%	0.005	0	0	75.0%	0	0.005	1/3	75.2%	0.01	0.005	1/3	**81.0%**
Nearest Neighbors	0	0	0	71.1%	0	0	0	72.2%	0.1	0	0	73.9%	0	0.01	1/3	74.9%	0.1	0.01	1/3	**77.0%**
Decision Tree	0	0	0	71.2%	0	0	0	72.4%	0.1	0	0	73.4%	0	0.01	1/3	75.8%	0.2	0.01	1/3	**77.3%**
LDA	0	0	0	76.5%	0	0	0	77.8%	0.05	0	0	79.8%	0	0.2	1/3	80.2%	0.2	0.2	1/3	**84.3%**
Linear SVM	0	0	0	79.6%	0	0	0	80.2%	0.005	0	0	81.6%	0	0.005	1/3	81.3%	0.2	0.005	1/3	**85.5%**
Logistic Regression	0	0	0	78.7%	0	0	0	79.7%	0.05	0	0	80.8%	0	0.2	1/3	81.8%	0.2	0.2	1/3	**85.3%**

TABLE III: Parameter impact on accuracy for MLP classifier

	λ_A	λ_N	r_N	**MLP**	**Adversary**	**Nuisance**
AE	0	0	0	72.2%	7.8%	5.6%
cAE	0	0	0	72.9%	8.5%	5.8%
D-cAE	**0**	**0.005**	**1/3**	**75.2%**	**8.6%**	**18.8%**
	0	0.01	1/3	74.1%	12.1%	24.2%
	0	0.2	1/3	72.3%	14.6%	35.5%
	0	0.5	1/3	74.9%	12.2%	47.5%
DA-cAE	**0.01**	**0.005**	**1/3**	**81.0%**	**6.1%**	**9.6%**
	0.1	0.005	1/3	78.0%	5.5%	9.7%
	0.2	0.005	1/3	80.3%	4.0%	11.1%
	0.5	0.005	1/3	78.5%	3.2%	14.0%

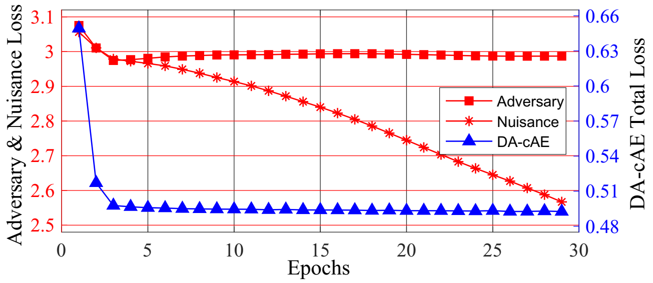


Fig. 3: Convergence of DA-cAE ($\lambda_A = 0.01$ and $\lambda_N = 0.005$).

of the nuisance unit declines steadily, while the adversary loss remains stable as a result of the adversarial relationship between the DA-cAE and adversary classifier, which keeps concealing subject-related information while not disabling the discriminative ability of overall network.

Finally, we evaluate the impact of training data size on the classification accuracy in Fig 4. Regardless of the data size reduction in the available training set, DA-cAE outperforms all models and shows robustness to physiological data size deficiency. Moreover, it is expected to achieve even higher gain when more measurement data are available for training.

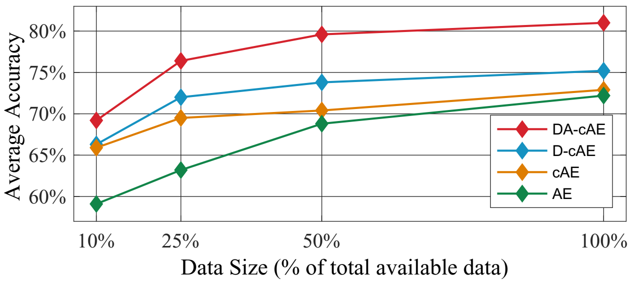


Fig. 4: Classification accuracies with different training dataset sizes, using the optimized model choices of Table III.

IV. Conclusion

We proposed a transfer learning method based on a disentangled adversarial AE model to extract nuisance-robust universal representations from physiological biosignals. To control the trade-off between task-related features and person-specific information, additional adversary and nuisance networks are jointly trained, such that the feature extractor is applicable to a broader range of users. We performed a cross-subject transfer analysis based on a public dataset for stress level assessment. Results showed remarkable benefits of the proposed method in improving both average and worst-case accuracies, indicating better adaptability to new subjects. Furthermore, extracted features showed universal robustness over different task-specific classifiers. We note that the proposed method is applicable to various other data analysis problems.

References

[1] S. Jerritta, M. Murugappan, R. Nagarajan, and K. Wan, "Physiological signals based human emotion recognition: a review," in *2011 IEEE 7th International Colloquium on Signal Processing and its Applications*. IEEE, 2011, pp. 410–415.

[2] P. C. Petrantonakis and L. J. Hadjileontiadis, "Emotion recognition from EEG using higher order crossings," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 2, pp. 186–197, 2009.

[3] M. Han, S. Y. Günay, G. Schirner, T. Padır, and D. Erdoğmuş, "HANDS: a multimodal dataset for modeling toward human grasp intent inference in prosthetic hands," *Intelligent Service Robotics*, vol. 13, no. 1, pp. 179–185, 2020.

[4] J. Birjandtalab, D. Cogan, M. B. Pouyan, and M. Nourani, "A non-EEG biosignals dataset for assessment and visualization of neurological status," in *IEEE International Workshop on Signal Processing Systems*, 2016, pp. 110–114.

[5] A. M. Amiri, M. Abtahi, A. Rabasco, M. Armey, and K. Mankodiya, "Emotional reactivity monitoring using electrodermal activity analysis in individuals with suicidal behaviors," in *10th International Symposium on Medical Information and Communication Technology*, 2016, pp. 1–5.

[6] D. Cogan, M. B. Pouyan, M. Nourani, and J. Harvey, "A wrist-worn biosensor system for assessment of neurological status," in *36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2014, pp. 5748–5751.

[7] D. Giakoumis, D. Tzovaras, and G. Hassapis, "Subject-dependent biosignal features for increased accuracy in psychological stress detection," *International Journal of Human-Computer Studies*, vol. 71, no. 4, pp. 425–439, 2013.

[8] G. Giannakakis, D. Grigoriadis, K. Giannakaki, O. Simantiraki, A. Roniotis, and M. Tsiknakis, "Review on psychological stress detection using biosignals," *IEEE Transactions on Affective Computing*, 2019.

[9] O. Özdenizci *et al.*, "Time-series prediction of proximal aggression onset in minimally-verbal youth with autism spectrum disorder using physiological biosignals," in *40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2018, pp. 5745–5748.

[10] H. Morioka, A. Kanemura, J.-i. Hirayama, M. Shikauchi, T. Ogawa, S. Ikeda, M. Kawanabe, and S. Ishii, "Learning a common dictionary for subject-transfer decoding with resting calibration," *NeuroImage*, vol. 111, pp. 167–178, 2015.

[11] W. Tu and S. Sun, "A subject transfer framework for EEG classification," *Neurocomputing*, vol. 82, pp. 109–116, 2012.

[12] S. Fazli, F. Popescu, M. Danóczy, B. Blankertz, K.-R. Müller, and C. Grozea, "Subject-independent mental state classification in single trials," *Neural Networks*, vol. 22, no. 9, pp. 1305–1312, 2009.

[13] O. Özdenizci, Y. Wang, T. Koike-Akino, and D. Erdoğmuş, "Learning invariant representations from EEG via adversarial inference," *IEEE Access*, vol. 8, pp. 27 074–27 085, 2020.

[14] ——, "Transfer learning in brain-computer interfaces with adversarial variational autoencoders," in *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*. IEEE, 2019, pp. 207–210.

[15] ——, "Adversarial deep learning in EEG biometrics," *IEEE Signal Processing Letters*, vol. 26, no. 5, pp. 710–714, 2019.

[16] G. Lample, N. Zeghidour, N. Usunier, A. Bordes, L. Denoyer, and M. Ranzato, "Fader networks: Manipulating images by sliding attributes," in *Advances in Neural Information Processing Systems*, 2017, pp. 5967–5976.

[17] M. Han, O. Özdenizci, Y. Wang, T. Koike-Akino, and D. Erdoğmuş, "Disentangled adversarial transfer learning for physiological biosignals," in *42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2020.

[18] F. Wu, X.-Y. Jing, Z. Wu, Y. Ji, X. Dong, X. Luo, Q. Huang, and R. Wang, "Modality-specific and shared generative adversarial network for cross-modal retrieval," *Pattern Recognition*, p. 107335, 2020.

[19] Y. Sun, X.-Y. Jing, F. Wu, J. Li, D. Xing, H. Chen, and Y. Sun, "Adversarial learning for cross-project semi-supervised defect prediction," *IEEE Access*, vol. 8, pp. 32 674–32 687, 2020.

[20] M. Atzori, M. Cognolato, and H. Müller, "Deep learning with convolutional neural networks applied to electromyography data: A resource for the classification of movements for prosthetic hands," *Frontiers in Neurorobotics*, vol. 10, p. 9, 2016.

[21] O. Faust, Y. Hagiwara, T. J. Hong, O. S. Lih, and U. R. Acharya, "Deep learning for healthcare applications based on physiological signals: A review," *Computer Methods and Programs in Biomedicine*, vol. 161, pp. 1–13, 2018.