Signal Processing and Machine Learning for Diplegia Classification

Luca Bergamini[1]([✉]), Simone Calderara[1], Nicola Bicocchi[1], Alberto Ferrari[2], and Giorgio Vitetta[1]

[1] Department of Engineering "Enzo Ferrari", University of Modena and Reggio Emilia, Via Vivarelli 10, 41125 Modena, Italy
179844@studenti.unimore.it
[2] Department of Electrical, Electronic and Information Engineering "Guglielmo Marconi", University of Bologna, viale Risorgimento 2, 40136 Bologna, Italy

Abstract. Diplegia is one of the most common forms of a broad family of motion disorders named *cerebral palsy* (CP) affecting the voluntary muscular system. In recent years, various classification criteria have been proposed for CP, to assist in diagnosis, clinical decision-making and communication. In this manuscript, we divide the spastic forms of CP into 4 other categories according to a previous classification criterion and propose a machine learning approach for automatically classifying patients. Training and validation of our approach are based on data about 200 patients acquired using 19 markers and high frequency VICON cameras in an Italian hospital. Our approach makes use of the latest deep learning techniques. More specifically, it involves a multi-layer perceptron network (MLP), combined with Fourier analysis. An encouraging classification performance is obtained for two of the four classes.

1 Introduction

Cerebral palsy (CP) is a broad family of non curable disorder of the voluntary muscular system, which appears in human's early childhood; this disorder is characterized by a great variety of symptoms, including stiff muscles, tremors and a general loss of coordination. CP is treatable but it is not curable; its symptoms could become more noticeable, but they do not worsen during lifetime. Common treatments include the effort of therapists and rehabilitation specialists, and the use of physiotherapy, antispastic drugs, orthosis and devices and functional surgery. Several classifications have been proposed for CP and are based on different aspects of this disorder. One of the most used classification relies on the identification of muscle tone anomalies as well on the type of the prevailing neurological symptom. This classification identified three classes as: (1) *spastic*, characterized by constant muscle tightness and stiffness; (2) *dyskinetic*, affecting patients unable to control involuntary movements; (3) *ataxic*, associated with shakiness and lack of coordination.

Another classification relies on the somatic location of the prevailing neurological symptom: (1) *tetraplegia* if all four limbs are affected; (2) *hemiplegia* if only one side of the body is afflicted; (3) *diplegia* if it involves symmetrical parts of the body.

A good classification system separates patients into clinical clusters characterized by sharing comparable prognosis, thus easing choice of treatment and communication of the expectations on autonomy level during child growth. With this goal, Ferrari *et al.* [7] have proposed a new classification of spastic forms of CP, which divides diplegia into other 4 forms (see Table 1), aimed at quickly conveying a "clinical snapshot" of a child by cross-referring to histories of other patients with similar motor impairments and dysfunctions therefore easing the choice of providing specific indications about the treatment to be adopted and about the disorder evolution over time. Moreover, it has been first validated in [4], analyzing a group of 467 subjects affected by CP (213 suffering from diplegia and 115 from tetraplegia), and characterized by significant correlations between identified walking forms. Further validation results, referring to the classification of spastic diplegia and involving 50 children and adolescents followed by professionals of rehabilitation, have been illustrated in [14]. This validation activities have evidenced that the less and the most severe forms of CP are the most easily identifiable, whereas the remaining two being more challenging.

In this work an automatic classification tool able to identify the 4 forms of spastic diplegia defined in [7] is illustrated. This tool combines frequency domain processing of the measurements acquired by means of multiple markers and high frequency VICON cameras with state of the art deep learning techniques. This work falls within the broad category of social signal processing, intended as the analysis of one ore more subjects as it interacts with another or with the environment. In this field techniques have shown competitive performance in classifying people interaction in crowd, as shown in [17]), but also in small groups [16] and in pairs [13].

The remaining part of this manuscript is organized as follows. In Section 2 previous work on the classification of diplegia is illustrated. In Sect. 3 gait analysis is introduced, and the sensors and methods employed for collecting measurements are illustrated; moreover, some indications about the adopted preprocessing techniques are provided. In Sect. 4 the architecture of the employed deep network is described. In Sect. 5 some numerical results are shown. Finally, Sect. 6 offers some conclusions.

2   Related Work

As far as we know, this is the first attempt of building a system based on the classification proposed in [7]. In the following previous work on diplegia classification is illustrated. However, note that most of it aims at discerning between healthy patients and patients affected by diplegia.

In [6] previous work on children affected by CP has been analysed. The considered methods include traditional data analysis systems and more recent machine learning techniques.

Table 1. Diplegia forms from Ferrari et al. [7]

Form	Traits
Form I-forward leaning propulsion	Antepulsion of trunk, walking on toe tips. Requires constant support of four point canes
Form II-tight skirt	Pronounced knee flexion in midstance, loaded knee behavior, short and frequent steps
Form III-tight rope walkers	Frontal trunk swinging and use of upper limbs to keep balance, presence of a dysperceptive disorder causing fearing feeling of falling and fear of open spaces without anything to cling on
Form IV-dare devils	Mainly a motor deficit. Increased talipes equinus at start of walking. Inability to stop walking abruptly without falling in early age

In [9] a *support vector machine* (SVM) for identifying spastic diplegia is proposed. A dataset of 3D points acquired by a six-camera Vicon System is analyzed; a group of 88 children affected by spastic diplegia and a control group of 68 children has been considered. Four features (namely, stride length, cadence, leg length and age) have been extracted from raw data. The best results have been achieved using only stride length and cadence (normalized on the basis of leg length and age) and adopting a radial basis function as kernel. In particular, an overall accuracy of 96.80% with 10-fold stratified cross validation has been achieved.

In [11] an *artificial neural network* is employed to combine traditional patient information with the analysis of heart rate variability. The proposed method is tested using a dataset that concerns healthy subjects and patients diagnosed with central coordination disturbances. Data have been acquired employing a 24-h ECG-Holter monitor; moreover, a shallow network consisting of a single hidden layer with 12 neurons has been employed.

In [1] a *Self-Organizing Map* (SOM) for unsupervised learning has been employed. The dataset collects information about three-dimensional joint angles, moments and powers and refers to 129 gait cycles from 18 subjects not affected by movement disorders; moreover, the *quantisation error* (QE) of the differences between normal and abnormal gaits is computed.

In [15] a rich dataset, referring to more than 900 patients affected with various pathologies, is exploited; data acquisition is based on a VICON 370 system consisting of high resolution infrared cameras. Moreover, the available data are processed to estimate hip rotation and the movements of other junctions, and then employed to train a SVM *classifier*. Since different disorders were included in the dataset, a binary classification algorithm for each of them against all the others has been trained.

Finally, the *principal component analysis* (PCA) has been employed in [3] to identify relevant information for classification of healthy and diplegic subjects.

3   Gait Analysis

Gait analysis plays a key role not only in the study of the muscular system, but also in that of the nervous and the sensory ones. Since this analysis involves the study of locomotion and of body mechanics, various devices are needed to simultaneously control the motion of multiple human joints; some of these devices are listed in Table 2. Gait analysis is usually accomplished in dedicated laboratories, called *motion analysis labs* (MALs).

Table 2. Devices commonly used for *gait analysis*

Devices	Description
Opto-electronic devices	Reflecting markers applied to the patient's skin in defined anatomical landmarks able to provide 3D motion of human segments
Force plates	Plates placed on the ground, used to measure the ground reaction force to body weight
Electromyography	Skin electrodes capable of acquiring the electrical signal generated by the contraction of muscles
Video systems	Cameras to record a patient during his/her trial

3.1   Data Format

One of the most advanced protocol for the storage of bio-mechanical data is the Total3DGait [2], which relies on the C3D format. A C3D file is composed by a *header* (which contains not only information about the considered patient such as his/her height or age, but also data useful for the remaining part of the file) and a *body*, which contains the position of every marker for each frame, along with the validity of the data itself, where the special value -1 indicates an invalid acquisition, and the precision of the measure. The order of the 3D positions is specified in the header. An insight in the single marker point is provided in Table 3.

3.2   Acquisition Method

As the gait analysis involves data referring to several distinct joints, a large number of markers needs to be applied to the skin of the considered patient. The *Davis protocol* [5] is one of the leading methodologies for acquiring not only 3D motion data, but also useful information about the considered patient, such as his/her weight and age; these data are collected before applying the opto-electronics markers for capturing kinematic, dynamic and electromyography measurements. Measurements are acquired as the patient walks at normal speed through the room hosting a MAL; the data usually refer to a variable number of trials (usually between 4 and 6).

Table 3. C3D frame

Word	Content (signed integer format)
1	X coordinate of point divided by POINT:SCALE factor
2	Y coordinate of point divided by POINT:SCALE factor
3	Z coordinate of point divided by POINT:SCALE factor
4	Byte 1: cameras measuring the considered marker (1 bit per camera). Byte 2: average residual divided by POINT:SCALE factor

3.3   Data Pre-processing

Our dataset refers to 1121 trials acquired from 178 patients affected by diplegia. The acquisition frame rate (100 frames/sec) provided by the employed VICON system is too large; for this reason every walk has been subsampled by a factor 2, so dropping the frame rate to 50 frames/sec. Moreover, trials not containing one or more of the 19 markers listed in Table 4 or trials referring less than two consecutive steps have been discarded. As the foot strikes were included as meta-data in every trial, multiple consecutive steps have been considered as a *sequence*. Our dataset evidences an heterogeneous distribution of the available measurements over the classes (see Table 6). This reflects, on the one hand, the incidence of different forms of this disorder in the population; on the other hand, it shows the difficulty for patients suffering from the most severe symptoms to sustain multiple trials. As it is fundamental to have a complete separation between the train set and the test set, our dataset has been split according to the proportion 0.75:0.25 patient-wise for each class. In Table 6 the resulting numbers are given for each class; note that the numbers indicated in bold refer to the case of data augmentation by repetition (this has been employed for the first class only, since it was substantially poorer than the remaining three classes).

Following [12], we have transformed the acquired measurements from the time domain to the frequency domain; in fact, this form of processing, implemented through the *fast Fourier transform* (FFT), has been shown to be extremely useful in discerning abnormal gaits from normal ones. Since the overall number N of steps performed by patients in their trials is highly variable among the four classes, only one coefficient every N was retained in the FFT output; this removes the dependence from the temporal length of the sequence. Moreover, only the first 20 coefficients selected in this way have been processed by our classification algorithm (apart from the first one, they have been normalized).

Moreover, before FFT processing, the following tasks have been accomplished:

1. The acquired xyz points have been projected onto two of the three human body's planes (in particular, the longitudinal and trasversal planes, with the third one corresponding to the floor plane). The use of these projections is fundamental to refer every trial to the same system of 3D coordinates. We also

Table 4. Markers: identifiers and positions

Name	Marker position	Name	Marker position
C7	c7 vertebra	LPSIS	Right hip joint
LA	Left shoulder	RGT	Right knee joint
RA	Right shoulder	LGT	Left knee joint
REP	Right elbow	RLE	Right ankle joint
LEP	Left elbow	LLE	Left ankle joint
RUL	Right wrist joint	RCA	Right heel joint
LUL	Left wrist joint	LCA	Left heel joint
RASIS	Right hip joint	RFM	Right foot
LASIS	Left hip joint	LFM	Left foot
RPSIS	Right hip joint		

swept two axes if the patient was to hindered to walk along the main side of the room hosting the MAL.

2. The absolute xyz coordinates have been transformed into a set of 27 three-dimensional angles, as shown in Table 5; each angle has been projected on every plane, as angles represent meaningful information in the classification of diplegia [7].

4   Classification Algorithm Based on a Multi Layer Perceptron Network

4.1   MLP Architecture

The base unit of a *multi layer perceptron* (MLP) network is called *perceptron*. Perceptrons have an high degree of similarity with the mammals brain cells, as they propagate or soften the incoming input from others. Stacking this units forms a layer, which could also be concatenated to others to obtain a network, where every layer is fully connected with the previous and the next one. To model the neuron activity on the incoming signal, a mathematical variable named weight is used for every connection, plus a bias is introduced to shift every output if it's needed. On the layer's output a non-linear function is then applied to map the layer's input to a new domain, which could be the network final output or the input of the next layer. This relation can be expressed as

$$y = f(W \circ x + b) \tag{1}$$

where x is the layer's input and the new output y, W and b are the weight matrix and the biases of the layer, respectively, $\circ$ is the dot product and f a continuous and differentiable function.

The initial values of the weight are drawn from a random distribution, which could be a Gaussian or some more complex models, while the bias are set to a

Table 5. Absolute xyz collected coordinates have been transformed into 27 three-dimensional angles, as most of the clinical signs of diplegia in all its forms are heavily related on angles instead of positions

Marker I	Marker II	Marker III
LGT	LPSIS	LLE
LLE	LGT	LCA
LCA	LLE	LFM
LEP	LA	LUL
LEP	C7	LUL
LLE	LASIS	LFM
LA	C7	LEP
RGT	RPSIS	RLE
RLE	RGT	RCA
RCA	RLE	RFM
REP	RA	RUL
REP	C7	RUL
RLE	RASIS	RFM
RA	C7	REP
LPSIS	LGT	RGT
LASIS	LGT	RGT
LPSIS	LLE	RLE
C7	LA	RA
C7	LEP	REP
RPSIS	LGT	RGT
RASIS	LGT	RGT
RPSIS	LLE	RLE
C7	LUL	RUL
LASIS	C7	LPSIS
RASIS	C7	RPSIS
LA	LASIS	RASIS
RA	LASIS	RASIS

small value (typically zero). To compare the output of the last network layer, which have a size equals to the classes number, with the labels, a mapping

$$y_{encode} = [\mathbf{0}\dots\mathbf{0},\mathbf{1},\mathbf{0}\dots\mathbf{0}] \tag{2}$$

$$with \begin{cases} y_{encode}[i] = 1 & if \quad i == class(y) \\ y_{encode}[i] = 0 & otherwise \end{cases} \tag{3}$$

called *one-hot encoding*, is required.

Table 6. Distribution of patients, trials and sequences over the four classes before and after data augmentation. The partitioning we adopted for training and test phases is also shown.

Class	Train		Test	
	Patients	Trials	Patients	Trials
0	9	47(**94**)	4	16
1	36	183	13	83
2	25	174	9	49
3	58	372	20	114

The base architecture of the MLP network we employed is composed by a single layer containing a number of perceptrons equal to the classes. We then added other layers to the first one; the number of hidden units included in each layer was twice that of the previous layer, starting from 32 units contained in the first additional layer (i.e., in the second layer of the network). To avoid a potential overfitting, a *dropout layer* [18] has been also employed; this randomly turns off some perceptrons during the training set, forcing the network not to rely on the same weights to produce a specific output.

The network providing the best results is represented in Fig. 1, (its accuracy scores are given in Sect. 5). Since a probabilistic interpretation of the output was required, a *softmax layer* was also used at the bottom of the network and Adam was used as optimizer [10]. In the following, unless differently stated, the accuracy metric refers to the single patient's trial only.
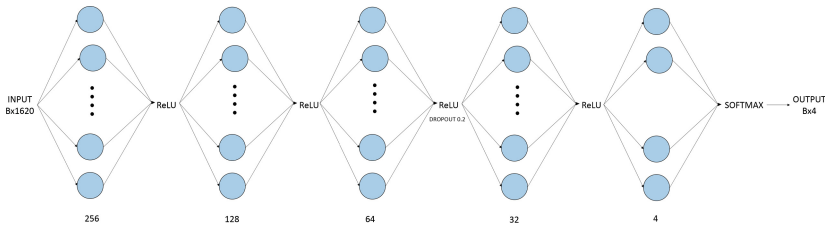

Fig. 1. Architecture of the proposed MLP network.

4.2 Training Phase

Using a part of the total dataset called train set, the MLP follows an iterative process of two steps. During the first one, called *forward propagation*, the train set is fed to the network until it reaches the last layer and the loss score is computed. In our work we used the *categorical cross entropy loss*

$$L = -\frac{1}{N * M} \sum_{n=0}^{N} \sum_{m=0}^{M} (y_{t[n,m]} * ln(y_{p[n,m]})) \qquad (4)$$

given the prediction $y_{p[n,m]}$ and the ground truth $y_{t[n,m]}$ for the n-th sample and the m-th class (M and N denote the overall number of classes and samples, respectively).

In the second step, called *backward propagation*, the loss score is used to update the network weights and the biases, in accordance with the gradient direction. The general algorithm, known as Stochastic Gradient Descent (SGD), for updating a weight w is shown in 5

$$w - \alpha * \frac{\nabla L}{\nabla w} \qquad (5)$$

where $\frac{\nabla L}{\nabla w}$ is the gradient $w.r.t.$ to w and α is a small constant called *learning rate.*

While the SGD is still widely used in the community, several optimized variants have been proposed in the past years [10].
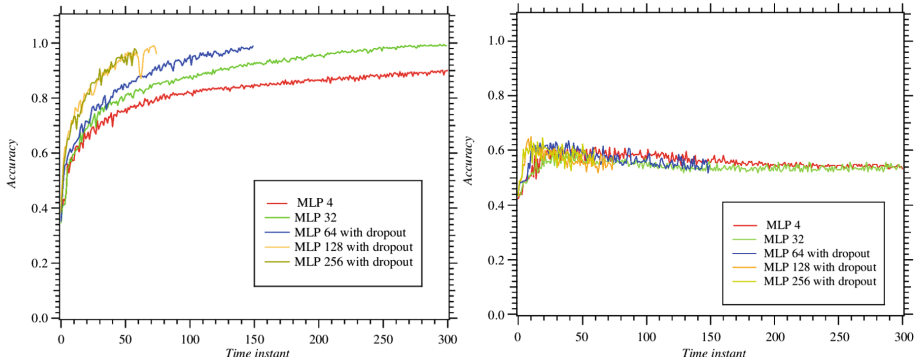

Fig. 2. Accuracies achieved by different MLP networks on both the training (left) and test (right); distinct networks are characterized by different numbers of hidden units in the most populated layer (see the labels of the curves).

5   Results

In Fig. 2 the trends of the accuracies achieved during both train and test phases for a set of MLP networks are shown. The networks are numbered with the hidden units of the most populated layer, and have been trained for an average of 500 epochs on an Nvidia GTX-1060. The network with 256 hidden units on the top layer has been the best performer, with about **0.603** on the test trials, while the other score lower results. While the accuracy reported refers to the single

trial, in Table 7 the class with the top frequency among the patient's trials is used as an accuracy meter, where every trial contributes to the final class prediction of the patient. The Accuracy of the MLP is tested against a baseline obtained using a Support Vector Machine Classifier [8] with radial basis kernel functions. While on the classes 0, 1 and 3 the results suggest a validation of the proposed classification system, the prediction on the class 2 are unreliable even with the top two scores using the MLP. The reason could lie in the main trait of this form being a perceptual disturb [7], which of course does not emerge from the motion data, remaining outside from the network knowledge.

Table 7. Accuracy scores (test set).

Class	MLP		SVM	
	Top one	Top two	Top one	Top two
0	0.75	0.75	0.0	0.15
1	0.846	1.0	0.461	0.615
2	0.111	0.333	0.0	0.777
3	0.6	0.9	0.9	0.95
Overall	0.595	**0.816**	0.537	0.761

The conclusion confirms the existence of the classification in accordance with the human's perception, since the classifier performs better on opposite classes which should have very different traits, as shown in Table 8. Among the other two classes, the number 2 is the most difficult to be classified, being often confused with the second and the fourth in the top one and top two predictions. In accordance with the experts in [14] this could be related to some forms of Diplegia being partially overlapped and thus having some common traits difficult to be discerned. The small size of the dataset makes difficult to validate several aspects of the learning process, such as the presence of overfitting on the train set.

Table 8. Confusion Matrix of the patients for MLP.

```
                      Predicted
                      3  1 0 0
       Ground Truth 0 11 2 0
                    0  4 1 4
                    0  7 1 12
```

6   Conclusion

In this work we proposed a method to tackle a four class problem using the state of the art deep learning techniques, which could aid in the development of more specific treatments for muscular system's pathologies, such as the Diplegia. We made use of a MLP (Multilayer Perceptron) on a dateset of 1121 trials from 178 patients, gathered in the last ten years at LAMBDA, Laboratorio Analisi del Movimento del Bambino Dis-Abile, Azienda Ospedaliera Arcispedale S. Maria Nuova and University of Modena and Reggio Emilia, Reggio Emilia, Italy. After a pre-processing step to extract Fourier coefficients of 3D angles motions, we fed these data to train a MLP able to identify 4 different Diplegia classes. Experimental results have been encouraging in 3 out of 4 classes. We make a commitment to release the full anonymized dataset to the machine learning community in future publications.

References

1. Barton, G., Lisboa, P., Lees, A., Attfield, S.: Gait quality assessment using self-organising artificial neural networks. Gait Posture **25**(3), 374–379 (2007)
2. Benedetti, M.G., Manca, M., Ferraresi, G., Cervigni, G., Berti, L., Leardini, A.: A new protocol for complete 3D kinematics analysis of the ankle foot complex in stroke patients. J. Foot Ankle Res. **1**(1), O30 (2008)
3. Carriero, A., Zavatsky, A., Stebbins, J., Theologis, T., Shefelbine, S.J.: Determination of gait patterns in children with spastic diplegic cerebral palsy using principal components. Gait Posture **29**(1), 71–75 (2009)
4. Cioni, G., Lodesani, M., Pascale, R., Coluccini, M., Sassi, S., Paolicelli, P.B., Perazza, S., Ferrari, A.: Eur. J. Phys. Rehabil. Med. **44**(2), 203–211 (2008)
5. Davis, R.B., Õunpuu, S., Tyburski, D., Gage, J.R.: A gait analysis data collection and reduction technique. Hum. Mov. Sci. **10**(5), 575–587 (1991)
6. Dobson, F., Morris, M.E., Baker, R., Graham, H.K.: Gait classification in children with cerebral palsy: a systematic review. Gait Posture **25**(1), 140–152 (2007)
7. Ferrari, A., Alboresi, S., Muzzini, S., Pascale, R., Perazza, S., Cioni, G.: The term diplegia should be enhanced. Part I: a new rehabilitation oriented classification of cerebral palsy. Eur. J. Phys. Rehabil. Med. **44**(2), 195–201 (2008)
8. Hearst, M.A., Dumais, S.T., Osuna, E., Platt, J., Scholkopf, B.: Support vector machines. IEEE Intell. Syst. Appl. **13**(4), 18–28 (1998)
9. Kamruzzaman, J., Begg, R.K.: Support vector machines and other pattern recognition approaches to the diagnosis of cerebral palsy gait. IEEE Trans. Biomed. Eng. **53**(12), 2479–2490 (2006)
10. Kingma, D., Ba, J.: Adam: a method for stochastic optimization. In: International Conference on Learning Representations, pp. 1–13 (2014)
11. Lukić, S., Ćojbašić, Ž., Jović, N., Popović, M., Bjelaković, B., Dimitrijević, L., Bjelaković, L.: Artificial neural networks based prediction of cerebral palsy in infants with central coordination disturbance. Early Hum. Dev. **88**(7), 547–553 (2012)
12. Mostayed, A., Mazumder, M.M.G., Kim, S., Park, S.J.: Abnormal gait detection using discrete fourier transform. In: International Conference on Multimedia and Ubiquitous Engineering, MUE 2008, pp. 36–40 (2008)

13. Palazzi, A., Calderara, S., Bicocchi, N., Vezzali, L., di Bernardo, G.A., Zambonelli, F., Cucchiara, R.: Spotting prejudice with nonverbal behaviours. In: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, pp. 853–862. ACM (2016)
14. Pascale, R., Perazza, S., Borelli, G., Bianchini, E., Alboresi, S., Paolicelli, P.B., Ferrari, A., Cioni, G.: The term diplegia should be enhanced. Part III: inter-observer reliability of the new rehabilitation oriented classification. Eur. J. Phys. Rehabil. Med. **44**, 213–220 (2008)
15. Salazar, A.J., De Castro, O.C., Bravo, R.J.: Novel approach for spastic hemiplegia classification through the use of support vector machines. In: 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEMBS 2004, vol. 1, pp. 466–469. IEEE (2004)
16. Setti, F., Russell, C., Bassetti, C., Cristani, M.: F-formation detection: individuating free-standing conversational groups in images. PLoS ONE **10**(5), e0123783 (2015)
17. Solera, F., Calderara, S., Cucchiara, R.: Socially constrained structural learning for groups detection in crowd. IEEE Trans. Pattern Anal. Mach. Intell. **38**(5), 995–1008 (2016)
18. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.: Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. **15**, 1929–1958 (2014)