Lower limb motor function assessment based on TensorFlow convolutional neural network and kernel entropy component analysis–local tangent space alignment

Yan Zhang¹,²,³, SiNing Li¹,²,³, Ying Zhou¹,² and Jian Liu¹,²,³

Abstract
Motor function assessment of patients and the elderly is crucial to gait assessment and gait rehabilitation. Accuracy of the assessment is affected by clinician’s experience. To solve the problem, this article proposes motor function assessment index to assess the motor function of patients. VICON system collects video of subjects when they are walking. And the original gait videos are pre-processed by the pixel-based adaptive segmenter and extracted by the convolutional neural network. The kernel entropy component analysis and local tangent space alignment reduced the dimensions of extracted features, and motor function assessment index is obtained. The Pearson correlation analysis shows that the motor function assessment index and modified gait abnormality rating scale are significantly correlated, and Pearson correlation coefficient is 0.92. These effectiveness results demonstrate that the proposed method has the considerable potential to promote the future design of automatic motor function assessment for clinical rehabilitation research.

Keywords
Convolutional neural network, lower limb motor function assessment, transfer learning, kernel entropy component analysis, local tangent space alignment

Introduction
The decline of lower limb motor function greatly reduces the elderly and the patients’ quality of life. Traditionally, the estimation of motor function is mostly based on the international classification of functioning, disability and health (ICF) and lower extremity motor score (LEMS). Nonetheless, these assessments are subjective and time-consuming, and their accuracy depends on the clinicians’ experience. Therefore, it is necessary to develop a convenient and objective quantitative assessment of motor function.

In the past few years, advanced measuring techniques have made it easier to collect the kinematic, kinetic, and gait video data during walking. It greatly facilitates the development of motor function assessment. Rozumalski and Schwartz⁵ and Schwartz and Rozumalski⁶ used three-dimensional gait information to evaluate lower limb motor function of cerebral palsy

¹School of Artificial Intelligence, Hebei University of Technology, Tianjin, China
²Engineering Research Center of Intelligent Rehabilitation, Ministry of Education, Hebei University of Technology, Tianjin, China
³Hebei Key Laboratory of Robot Sensing and Human-Robot Interaction, Hebei University of Technology, Tianjin, China

Corresponding author:
Ying Zhou, School of Artificial Intelligence, Hebei University of Technology, Tianjin 300310, China.
Email: zhouying2007@163.com
patients and established quantitative evaluation indexes. Gouelle et al. collected eight gait spatiotemporal parameters from patients with Friedrich's ataxia and developed gait variability index (GVI) to assess motor function. Ma and Liao collected ground reaction forces and presented abnormal evaluation index (AEI) which used a semi-Markov process. Rahmati et al. used lower limb surface electromyographic (sEMG) signals of cerebral palsy patients to establish synergy comprehensive assessment (SCA) for quantitative evaluation motor dysfunction. Modified gait abnormality rating scale (GARS-M) is a semi-subjective assessment method based on video, which evaluates seven gait characteristics during human walking, including general performance; lower limb movements; and the trunk, head, and upper limb movements. However, these assessment were usually calculated based on one kind of data source (e.g. kinematic data), which could only summarize a subject’s gait characteristics from a single aspect. And it might lead to a one-sided assessment result. Some researchers built assessment models based on multisource data. Soran et al. collected natural behavior videos of children with spinal muscular atrophy (SMA) and used regression convolutional neural network (RCNN) to analyze their behavioral videos. D’Souza et al. established a quantitative multiple sclerosis assessment system using deep video motion data, and the effectiveness of the system is verified by clinical trials.

In recent years, CNN has achieved great results in medical image diagnosis, video target detection, and behavior recognition areas. Deep convolution layers extracted more abstract features and got more effective recognition. Many previous studies showed that CNN had good feature extraction capabilities and revealed hidden and complicated features in the spatial domain. To avoid subjective effects, TensorFlow convolutional neural network (TF-CNN) is used to extract features from gait video. Kernel entropy component analysis (KECA)–local tangent space alignment (LTSA) reduces redundant features and got motor function assessment index (MFAI). Compared with traditional CNN, TF-CNN reduces the training time of CNN and improves the accuracy. KECA–LTSA is more efficient than linear discriminate analysis (LDA) and principal component analysis (PCA) in nonlinear data extraction.

The article is structured as follows. Section “Materials and methods” introduces the process of data collection and describes the new assessment based on improved CNN and KECA–LTSA. Section “Experimental results” shows the experiment results of proposed algorithm. Conclusion and future work are reported in section “Conclusion and future work.”

Materials and methods

Overview of the proposed method

As shown in Figure 1, this assessment system consists of data collection, data feature extraction, and motor function assessment. In the first stage, VICON system collects video data. Then, pixel-based adaptive segmenter (PBAS) extracts the silhouette images from video and TF-CNN extracts feature from images. KECA–LTSA reduced redundant features and MFAI is proposed.

Classification of subjects. All subjects of this article are divided into three groups by clinician, including middle-age stroke group (MG, 25 subjects), elderly stroke group (EG, 25 subjects), and young healthy group (YG, 40 subjects). GARS-M is used to assess the motor function of all subjects preliminary by clinicians. The GARS-M scores range from 0 to 21. Table 1 lists the conditions of the three group subjects, including age, gender, physical condition, and the GARS-M score. The MG group consists of subjects with motor dysfunction aged between 40 and 59 years with a GARS-M score of 3–9. The EG group consists of subjects with motor dysfunction aged between 60 and 70 years with a GARS-M score of 8–14. Both the EG and MG groups are the test group in this study. YG group is considered as reference group. The YG group consists of subjects with motor dysfunction aged between 20 and 39 years. Because there is no GARS-M score in young healthy group, the GARS-M score in YG group is 0. All the subjects could complete the data collection without any help or assistance.

Motor data collection. VICON Motion Capture System (by Oxford Metrics Limited, UK) is used to collected gait video. It consists of six cameras, signal capture

![Figure 1. Overview of the research scheme.](image)
equipment, and data generation software. Full image measurement frequency of cameras is 1000 Hz. In this study, the measurement frequency is conducted at 200 Hz. As shown in Figure 2, VICON captures the video of subjects when they are walking, while the silhouette images of video are shown in Figure 3.

**Data feature extraction**

As shown in Figure 4, feature extraction could be summarized as three steps

1. Subjects’ images are divided into train set and test set, whose ratio is 8–2.
2. CNN model parameters are pre-trained on the ImageNet and used as the initialization parameters of the network. The trained CNN model extracts features from gait silhouettes images.
3. The gait features are fused serially to obtain feature-vector.

**Structure of CNN.** The trained CNN is used as a feature extractor to establish the motor function assessment model. VGG-16 has a simple structure but performed a high accuracy on ImageNet. And transfer learning and fine-tuning show strong generalization ability on VGG-16 from big data set to small data set. As shown in Figure 5, CNN structure is modified from VGG-16. This network has seven sequentially connected layers, including input layer (input image size is 120 × 150 pixel), five convolution blocks, two full connection layers, and a output layer. Each convolution block consists of two or three convolution layers and one max-pooling layer. The output of Block5 is flattened into

---

**Table 1.** Demographic data of the subjects.

| Subjects       | Number | Age  | Physical condition | GARS-M | Height | Weight | BMI     |
|----------------|--------|------|--------------------|--------|--------|--------|---------|
| Reference group| YG     | 40   | 29.2 ± 6.2         | Young healthy | 0      | 173 ± 10 | 70 ± 8  | 22.5 ± 3.6 |
| Test group     | MG     | 25   | 48.2 ± 5.4         | Middle-age stroke | 6 ± 3  | 169 ± 6  | 55 ± 7  | 20.3 ± 4.3  |
|                | EG     | 25   | 63.8 ± 3.6         | Elderly stroke group | 11 ± 3 | 170 ± 8  | 56 ± 4  | 19.4 ± 2.7  |

GARS-M: modified gait abnormality rating scale; BMI: body mass index; YG: young healthy group; MG: middle-age stroke group; EG: elderly stroke group.
full connection layer vector. FC1 connection layer is a 256-dimensional vector and FC2 connection layer is used to gait recognition by softmax function.

Table 2 shows the structure of CNN. C represents convolution layer, R represents nonlinear activation function (RELU), P represents maximum pooling layer, F represents full connection layer, D represents Dropout (probability is 0.5), and S represents activation function. And the number of trainable parameters at each layer is given in Table 2.

Transfer learning and fine-tuning. Transfer learning is different from the traditional CNN training method. It could improve the performance of learning by transferring knowledge from several source domains to a target domain.22

In this study, CNN structure is trained by ImageNet data set, which includes one million pictures, to get parameters. Then, the pre-trained parameters are used as the initial parameters of Block1–Block5 layers, and the parameters of FC6 and FC7 are randomly initialized. During the training, the gait image data set is used to fine-tune and train the parameters of FC6 and FC7. Although there is a big difference between the images of ImageNet database and the gait video image of this study, pre-training weight is better than the random weight because the initialization of random weight has a high chance to fit the training data.

Motor function assessment

Feature normalization. The Gaussian feature normalization is used to process the features, so that the indicators of different magnitudes can be uniformly reduced to remove the correlation, and the local influence of different types of features is guaranteed. Combining the extracted features into a fusion matrix \( G = [g_1, \ldots, g_m, \ldots, g_N]_{m \times N} \), which \( m \) is the characteristic dimension of the fusion, \( N \) is the number of experimental objects, \( g_i \) is the fusion feature vector of the first experimental object, and the average value \( \mu_i \) and
where

\[
\hat{X}_n = U_n \Pi_n V_n^T = U_n \Omega_n = [U_{n,1} \quad U_{n,2}] [\Omega_{n,1}^e \quad \Omega_{n,2}^e]
\]

where \( \Omega_n = \Pi_n V_n^T \), the first \( d \) column vectors of \( U_n \) compose \( U_{n,1} \), and \( \Omega_{n,1} \) are the eigenvectors of \( U_{n,1} \). And \( T \) is defined as

\[
T = \lambda_{\text{max}} I - \Omega_{n,1}
\]

where \( \lambda_{\text{max}} \) is the maximum value of \( \lambda \) and \( I \) is the unit matrix. Kernel matrix \( K_{\text{LTSA}} \) is

\[
K_{\text{LTSA}} = (I - \frac{1}{N} ee^T) T \left( I - \frac{1}{N} ee^T \right)^T
\]

Therefore, the final kernel matrix is

\[
K_{\text{KECA-LTSA}} = (1 - \alpha) K_{\text{KECA}} + \alpha K_{\text{LTSA}}
\]

where \( \alpha \) is the balance parameter, \( K_{\text{KECA-LTSA}} \) can be represented by kernel matrix eigenvalues \( \lambda_d = [\lambda_1, \lambda_2, ..., \lambda_d] \) and eigenvectors \( \nu_d = [\nu_1, \nu_2, ..., \nu_d] \).

**MFAL**. The YG group’s average features in the subspace are defined as standard healthy gait. Every subject’s features are compared with standard features, the gait performance of the subject is obtained, and the lower limb MFAL can be evaluated.

The average of YG is

\[
\bar{C}_{\text{YG}}^{\text{Reference}} = \frac{1}{n} \sum_{i=1}^{n} t^{\text{YG}(i)}
\]
The deviation between other subjects and standard features is
\[ d^{\text{reference}} = \| t^{a} - \bar{c}_{YG}^{a} \| \quad (12) \]

The MFAI of all subjects is obtained as follows
\[ MFAI^{\alpha}_{\text{reference}} = \ln(d^{\alpha}_{\text{reference}}) \quad (13) \]

The gait deviation index of any of the subjects in the reference group YG is
\[ MFAI^{YG}_{\text{reference}} = \ln(d^{\beta}_{\text{reference}}) \quad (14) \]

The MEAN(Z) and S.D.(Z) represent the average of Z group and standard deviation of Z group. Zscore definition is
\[ Z\text{score} = \frac{Z - \text{Mean}(Z)}{\text{S.D.}(Z)} \quad (15) \]

The MFAI of all the subjects in the experiment is obtained as
\[ zMFAI^{\alpha}_{\text{reference}} = \frac{MFAI^{\alpha}_{\text{reference}} - \text{Mean}(MFAI^{YG}_{\text{reference}})}{\text{S.D.}(MFAI^{YG}_{\text{reference}})} \quad (16) \]

And a percentage system evaluation index is
\[ MFAI^{\alpha} = 100 - 10 \times zMFAI^{\alpha}_{\text{reference}} \quad (17) \]

**Experimental results**

**CNN analysis**

The training results of CNN using transfer learning are shown in Figure 6. In the figure, acc and loss represent the training set accuracy rate and the loss value, and val_acc and val_loss represent the accuracy and loss values of the test set. The results of TF-CNN are shown in Figure 7. It shows that transfer learning and fine-tuning training CNN converge faster than transfer learning CNN.

As shown in Table 3, TF-CNN is better than CNN. And after fine-tuned, TF-CNN’s accuracy is improved. Therefore, this article uses transfer learning and fine-tuning to train convolutional networks in the network training process.

**Feature extraction analysis**

The multi-feature fusion method used in this article avoided the randomness of single feature parameters, but the feature parameters increases exponentially. To solve this problem, the redundant information of the fusion matrix is removed by KECA–LTSA. Radial basis function (RBF) kernel function is used as the kernel function of KECA and the kernel width parameter \( \delta = 0.3 \), and balance parameter \( \alpha = 0.7 \). Kernel function of kernel principal component analysis (KPCA) is also the RBF kernel function and the kernel width parameter \( \delta = 0.5 \).

It can be seen that when the EC is 97%, the principal subspace is 30 dimensions. And when the variance contribution (VC) is 97%, it is 45 dimensions. So KECA–LTSA can cover the original data information with fewer feature dimensions. Table 4 shows the eigenvalues and cumulative EC ratios of the first 10 principal
components of KECA–LTSA, the corresponding eigenvalues of the first 10 kernel principal components of KPCA, and their cumulative VC ratios. It shows that the features extracted by KECA–LTSA are more interpretable than KPCA.

**MFAI indicators and statistical analysis**

Using the first 30 elements extracted by KECA–LTSA as the final feature vector, the subspace of the feature extraction is constructed, and the lower limb dyskinesia evaluation index (MFAI) is obtained.

The MFAI scores of each group extracted by KPCA are shown in Figure 8. YG scores range from 95 to 123. MG scores range from 95 to 83. EG scores range from 94 to 87. It can be seen that YG group can be distinguished from MG and EG groups, but it is difficult to classify MG and EG groups.

The MFAI scores extracted by KECA are shown in Figure 9. YG scores range from 90 to 116. MG scores range from 50 to 90. EG scores range from 43 to 66. KECA performance is better than KPCA. MG and EG groups’ distance is increased.

The MFAI scores extracted by KECA–LTSA are shown in Figure 10. YG group’s MFAI scores range from 91 to 112. MG group’s MFAI scores range from 86 to 72. EG group’s MFAI scores range from 61 to 40. This method has the best classification, and there is large distance between different groups and small distance within group.

**Correlation analysis between MFAI index and GARS-M**

GARS-M score of subjects with dysfunction (i.e. MG and EG groups), which evaluated by physicians, and

| Number | KPCA | KECA–LTSA |
|--------|------|-----------|
|        | Eigenvalue | VC (%) | Eigenvalue | EC (%) |
| 1      | 4.1132 | 12.90 | 7.3610 | 16.14 |
| 2      | 3.2641 | 23.14 | 4.2612 | 27.34 |
| 3      | 2.1714 | 29.96 | 5.1101 | 36.69 |
| 4      | 1.7158 | 35.34 | 4.1885 | 45.72 |
| 5      | 1.2687 | 39.32 | 3.1812 | 53.59 |
| 6      | 0.9035 | 42.16 | 3.5883 | 60.57 |
| 7      | 0.8672 | 44.88 | 1.9019 | 64.74 |
| 8      | 0.8325 | 47.49 | 1.3673 | 67.74 |
| 9      | 0.7816 | 49.95 | 1.2952 | 70.58 |
| 10     | 0.7083 | 52.17 | 0.7668 | 72.26 |

KPCA: kernel principal component analysis; KECA: kernel entropy component analysis; LTSA: local tangent space alignment; VC: variance contribution; EC: entropy contribution.

Figure 8. MFAI scores extracted by KPCA.

Figure 9. MFAI scores extracted by KECA.

Figure 10. MFAI scores extracted by KECA–LTSA.
proposed MFAI are analyzed with Pearson correlation coefficient. Figure 11 shows the linear regression analysis of the MFAI scores extracted by KECA–LTSA and GARS-M. The Pearson correlation coefficient is $r = 0.92$. It can verify that the effectiveness of the MFAI.

Conclusion and future work

This article proposes a novel lower limb motor function assessment. First, VICON system collects subjects’ gait video data. CNN is pre-trained by ImageNet and extracts features from gait images. Compared with normal training method, transfer learning and fine-tuning could make CNN converge faster and have higher accuracy. Then, feature data are fused into feature matrix. KECA–LTSA is used to decrease the dimensions of the feature matrix. Compared with the KPCA and KECA, KECA–LTSA decreases more dimensions and retains effective information of original data, and the MFAI results shows that classification of KECA–LTSA is better than KPCA and KECA. And the correlation coefficient between MFAI and GARS-M scores is 0.92, which proves the effectiveness of the method.

In this article, we only studies how to evaluate dysfunction of lower limb, which we think affects human’s motion function directly. However, some diseases and dysfunction of other parts of body could affect lower limb’s function, such as diabetes and obesity. In the future work, this research can be extended to evaluate more people who have complicated healthy conditions to further improve the model.

Author contributions

Y.Z. was involved in conceptualization, methodology, and writing—review and editing. S.L. was involved in data curation, investigation, and writing—original draft. Y.Z. was involved in writing—review and editing. J.L. was involved in data curation and writing—review and editing.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by National Science Foundation of China (Grant Nos. 61773151 and 61703135) and Natural Science Foundation of Hebei Province (Grant No. F2018202279).

ORCID iD

SiNing Li https://orcid.org/0000-0001-5514-2821

References

1. Blanpied PR, Gross AR, Elliott JM, et al. Neck pain: revision 2017. J Orthop Sport Phys 2017; 47: A1–A83.
2. Werle J and Hauer K. Design of a bath robot system—user definition and user requirements based on International Classification of Functioning, Disability and Health (ICF). In: Proceedings of the 2016 25th IEEE international symposium on robot and human interactive communication (RO-MAN), New York, 26–31 August 2016. New York: IEEE.
3. Parijat P, Lockhart TE and Liu J. EMG and kinematic responses to unexpected slips after slip training in virtual reality. IEEE Trans Biomed Eng 2015; 62: 593–599.
4. Terrier P, Le Carre J, Connaissa ML, et al. Monitoring of gait quality in patients with chronic pain of lower limbs. IEEE Trans Neural Syst Rehabil Eng 2017; 25: 1843–1852.
5. Rozumalski A and Schwartz MH. The GDI-kinetic: a new index for quantifying kinetic deviations from normal gait. Gait Posture 2011; 33: 730–732.
6. Schwartz MH and Rozumalski A. The gait deviation index: a new comprehensive index of gait pathology. Gait Posture 2008; 28: 351–357.
7. Gouelle A, Mégrot F, Presedo A, et al. The gait variability index: a new way to quantify fluctuation magnitude of spatiotemporal parameters during gait. Gait Posture 2013; 38; 461–465.
8. Ma H and Liao WH. Human gait modeling and analysis using a semi-Markov process with ground reaction forces. IEEE Trans Neural Syst Rehabil Eng 2016; 25: 597–607.
9. Rahmati H, Martens H, Aamo OM, et al. Frequency analysis and feature reduction method for prediction of cerebral palsy in young infants. IEEE Trans Neural Syst Rehabil Eng 2016; 24: 1225–1234.
10. Vandenberg JM, George DR, O’Leary AJ, et al. The modified gait abnormality rating scale in patients with a conversion disorder: a reliability and responsiveness study. Gait Posture 2014; 95: e98–e98.
11. Soran B, Lowes L and Steele KM. Evaluation of infants with spinal muscular atrophy type-I using convolutional
neural networks. In: Proceedings of the European conference on computer vision, Amsterdam, 8–16 October 2016. Cham: Springer.

12. D’Souza M, Burggraaff J, Steinheimer S, et al. Accuracy of depth-sensing recordings in classifying expanded disability status scale subscores of motor dysfunction in patients with multiple sclerosis. Am Acad Neurol Poster Present 2016; 86: 132–139.

13. Liu R, Zhao Y, Wei S, et al. Indexing of CNN features for large scale image search. Pattern Recogn 2015; 48: 2983–2992.

14. Karpathy A, Toderici G, Shetty S, et al. Large-scale video classification with convolutional neural networks. In: Proceedings of the 2014 IEEE conference on computer vision and pattern recognition (CVPR), Columbus, OH, 24–27 June 2014. New York: IEEE.

15. Kido S, Hirano Y and Hashimoto N. Detection and classification of lung abnormalities by use of convolutional neural network (CNN) and regions with CNN features (R-CNN). In: Proceedings of the 2018 international workshop on advanced image technology (IWAIT), Chiang Mai, Thailand, 7–9 January 2018. New York: IEEE.

16. Wang X, Lu L, Shin HC, et al. Unsupervised joint mining of deep features and image labels for large-scale radiology image categorization and scene recognition. In: Proceedings of the 2017 IEEE winter conference on applications of computer vision (WACV), Santa Rosa, CA, 24–31 March 2017. New York: IEEE.

17. Batchuluun G, Yoon HS, Kang JK, et al. Gait-based human identification by combining shallow convolutional neural network-stacked long short-term memory and deep convolutional neural network. IEEE Access 2018; 6: 63164–63186.

18. Wang W and Shen J. Deep visual attention prediction. IEEE Trans Image Process 2017; 27: 2368–2378.

19. Senhadji L, Shu H, Zeng R, et al. Quaternion softmax classifier. Electron Lett 2014; 50: 1929–1931.

20. Liu B, Yu X, Zhang P, et al. Supervised deep feature extraction for hyperspectral image classification. IEEE Trans Geosci Remote Sens 2018; 56: 1909–1921.

21. Xu X, Li W, Ran Q, et al. Multisource remote sensing data classification based on convolutional neural network. IEEE Trans Geosci Remote Sens 2017; 56: 937–949.

22. Chen C, Jiang B and Jin X. Parameter transfer extreme learning machine based on projective model. In: Proceedings of the 2018 international joint conference on neural networks (IJCNN), Rio de Janeiro, 8–13 July 2018, pp.1–8. New York: IEEE.