

تشخیص حرکات مچ دست از روی سیگنال الکترومایوگرام با استفاده از شبکه عصبی کانولوشنال

چکیده

دریافت: ۱۳۹۸/۰۵/۱۶ ویرایش: ۱۳۹۸/۰۵/۲۳ پذیرش: ۱۳۹۸/۰۷/۲۰ آتالین: ۱۳۹۸/۰۷/۳۰

علی عامری*

گروه مهندسی پزشکی، دانشکده پزشکی،
دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران.

زمینه و هدف: با پیشرفت یادگیری عمیق (Deep learning)، انقلاب بزرگی در هوش مصنوعی ایجاد شده که بسیاری از رشته‌ها را به شدت تحت تاثیر خود قرار داده است. یادگیری عمیق، پردازش داده‌های خام با ابعاد بالا (مانند سیگنال یا تصویر) را بدون نیاز به مهندسی ویژگی (Feature engineering)، امکان‌پذیر می‌کند. هدف از این پژوهش، توسعه یک سیستم بر پایه یادگیری عمیق، برای تخمین اراده حرکتی از روی سیگنال EMG می‌باشد.

روش بررسی: در این مطالعه، یک سیستم مایوالکتریک (Myoelectric) بر پایه شبکه عصبی کانولوشنال (CNN) (که یک مدل یادگیری عمیق است)، به‌عنوان جایگزینی برای روش‌های معمول طبقه‌بندی (Classification) که نیازمند به مهندسی ویژگی هستند، معرفی شده است. این سیستم برای حرکات انفرادی و ترکیبی مچ دست، بر روی ده شخص سالم، مورد ارزیابی قرار گرفته شد. عملکرد روش پیشنهادی، با یک سیستم استاندارد برپایه Support vector machine (SVM) که از ویژگی‌های حوزه زمانی (Time domain, TD) استفاده می‌کند، مقایسه گردید.

یافته‌ها: با وجود عملکرد ثابت شده و رواج بسیار بالای ویژگی‌های TD، سیستم پیشنهادی به‌دقت طبقه‌بندی مشابهی ($P=0/19$) دست یافت. مزیت سیستم پیشنهادی در این است که نیازی به استخراج دستی و مهندسی ویژگی از سیگنال EMG وجود ندارد و CNN به‌صورت خودکار، ویژگی‌های مورد نیاز را فراگرفته و از سیگنال استخراج می‌کند.

نتیجه‌گیری: این یافته‌ها، توانایی بالای CNN، برای یادگیری و استخراج اطلاعات غنی و پیچیده از سیگنال‌های بیولوژیک را نشان می‌دهد. CNN می‌تواند اطلاعات زمانی و فرکانسی مورد نیاز برای تخمین اراده حرکتی را از روی سیگنال EMG فرا بگیرد.

کلمات کلیدی: طبقه‌بندی، شبکه عصبی کانولوشنال، یادگیری عمیق، الکترومایوگرام.

* نویسنده مسئول: تهران، ولنجک، دانشگاه علوم
پزشکی شهید بهشتی، دانشکده پزشکی.

تلفن: ۰۲۱-۲۲۴۳۹۹۴۱

E-mail: ameri@sbmu.ac.ir

مقدمه

یکی از کاربردهای سیستم‌های مایوالکتریک، کنترل دست یا پای مصنوعی متحرک است.^۱ در این سیستم‌ها، حس‌گرهای سطحی بر روی پوست، سیگنال‌های EMG را ضبط کرده و از روی EMG به کمک روش‌های یادگیری ماشین (Machine learning)، اراده حرکتی (Motor intent) تخمین زده می‌شود. این اراده حرکتی تخمین زده‌شده، توسط سیگنال کنترلی مناسب برای حرکت دادن اندام مصنوعی به آن

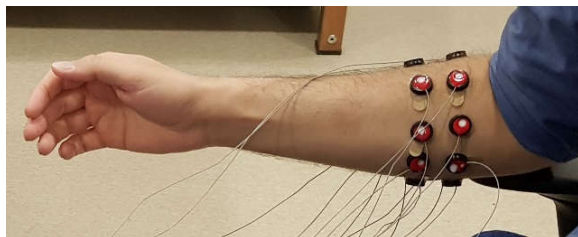
ارسال می‌شود.^۱ یکی از روش‌های رایج یادگیری ماشین در سیستم‌های مایوالکتریک، طبقه‌بندی (Classification) می‌باشد.^۲ به‌دلیل اینکه سیگنال EMG یک سیگنال استوکستیک (رندم) است، مقادیر لحظه‌ای آن برای تخمین اراده حرکتی مناسب نیستند.^۳ به‌منظور حل این مشکل، یک مجموعه از ویژگی‌ها از پنجره‌های زمانی (به‌طور معمول بین ۱۰۰ تا ۲۰۰ میلی‌ثانیه) سیگنال استخراج می‌شود.^۳ این ویژگی‌ها به‌عنوان ورودی برای سیستم طبقه‌بندی مورد استفاده قرار می‌گیرند. یکی از معروف‌ترین مجموعه ویژگی‌ها، مجموعه TD است که شامل چهار ویژگی میانگین

دارند، چراکه حرکتی طبیعی شبیه حرکت دست واقعی انسان را امکان پذیر می کنند.

روش بررسی

ده شخص سالم با سن $31/4 \pm 1/1$ ، ۹ راست دست، ۱ چپ دست در این مطالعه شرکت کردند. پروتکل آزمایش، توسط کمیته اخلاق دانشکده پزشکی دانشگاه علوم پزشکی شهید بهشتی مورد تایید قرار گرفت. همه سابجکت ها پیش از آزمایش، فرم رضایت نامه کتبی را امضا کردند. هشت جفت الکتروود سطحی (gHiAmp, gtec Inc.) به صورت هشت کانال دوقطبی برای ضبط EMG، به دور ساعد غالب، پایین تر از آرنج، با فواصل مساوی بر روی پوست چسبانده شدند (شکل ۱).

یک الکتروود مرجع هم بر روی Ulnar-styloid دست مقابل چسبانده شد. هشت حرکت مچ شامل فلکشن، اکستنشن، پرونیشن، سوپینیشن، و ترکیب آن ها: فلکشن-پرونیشن، فلکشن-سوپینیشن، اکستنشن-پرونیشن، و اکستنشن-سوپینیشن، مورد مطالعه قرار گرفتند. مطالعه بر روی حرکات دینامیک انجام شد. آزمایش شامل ۹ تریال بود که هر تریال مختص یکی از انقباض های ۸ گانه و یک تریال هم حالت استراحت بود. تریال استراحت ۳۰ ثانیه بود که کاربر عضلات خود را در حالت بدون انقباض نگه داشت. در ۸ تریال دیگر، کاربر انقباض مختص آن تریال را به صورت دینامیک، با دنبال کردن یک هدف دیداری بر روی یک صفحه کامپیوتر که پیش روی کاربر قرار داشت، انجام داد (شکل ۲). این هدف دیداری، طول یک خط



شکل ۱: چسبانده شدن هشت جفت الکتروود سطحی EMG به صورت ۸ کانال دوقطبی، به دور ساعد غالب

قدرمطلق، تعداد گذر از صفر، تعداد تغییر شیب منحنی و طول شکل موج است.^۴ مجموعه TD حاوی اطلاعات زمانی و فرکانسی سیگنال می باشد. برخی از ویژگی های پرکاربرد دیگر عبارتند از فرکانس میانگین، ضرایب اتورگرسیون، دامنه ویلیسون، توان میانگین و هیستوگرام.^۵

به فرآیند طراحی و استخراج ویژگی ها، مهندسی ویژگی گفته می شود. از آنجا که مهندسی ویژگی تاثیر زیادی بر دقت تخمین اراده حرکتی دارد، پژوهش های بسیاری بر روی آن انجام شده است.^۶ بسیاری از ویژگی های سیگنال EMG که در مطالعات پیشین^۵ معرفی شده اند، همبستگی (Correlation) زیادی با هم دارند، که سبب افزونگی (Redundancy) ویژگی می شود.^۷ بنابراین، به دست آوردن ویژگی هایی که غیرزاید هستند و با کمترین پیچیدگی محاسباتی منجر به جداسازی بهینه کلاس ها در سیستم طبقه بندی می شوند، از اهمیت بالایی برخوردار است. بسته به کاربرد، این ویژگی ها باید قابل تعمیم بین مجموعه داده های متنوع باشند.^۷ با توجه به این اهمیت، سیستمی با توانایی شناسایی و استخراج خودکار ویژگی های مفید سیگنال EMG، بسیار ارزشمند خواهد بود. روش های یادگیری عمیق، این توانایی را دارا می باشند. در سال های اخیر، یادگیری عمیق، رشته های بسیاری از جمله بینایی ماشین، تشخیص گفتار، و بیوانفورماتیک را متحول کرده است و به عملکردهایی مشابه یا بالاتر از عملکرد انسان دست یافته است.^۸

شبکه عصبی کانولوشنال (Convolutional neural network, CNN) یک روش یادگیری یاد عمیق است. CNN مانند دیگر روش های یادگیری عمیق، می تواند طبقه بندی را به طور مستقیم از روی داده خام بدون نیاز به مهندسی ویژگی انجام دهد. در واقع CNN می تواند ویژگی های مورد نیاز طبقه بندی را از روی داده خام یاد گرفته و استخراج کند. یکی از دلایل افزایش روزافزون کاربرد CNN، پیشرفت های اخیر در GPUها (پردازنده های گرافیکی) است که زمان آموزش شبکه را به میزان چشمگیری کاهش داده اند. در سه سال اخیر در چند مطالعه، CNN برای تشخیص اراده حرکتی از روی EMG به کار رفته است.^{۹-۱۲}

در این مقاله، یک شبکه CNN جدید برای تشخیص حرکات ترکیبی و انفرادی مچ از روی EMG پیشنهاد شده است. حرکات ترکیبی در مطالعات پیشین که از CNN استفاده کرده اند، مورد بررسی قرار نگرفته اند،^{۹-۱۲} اما اهمیت بسزایی در کنترل دست های مصنوعی

دارد که شامل یک لایه ورودی و پنج لایه کانولوشن است. پس از هر لایه کانولوشن، یک لایه Batch-normalization، یک لایه Relu، و یک لایه Max-pooling (برای سه لایه کانولوشن اول) به کار گرفته شده است. در آخر هم لایه Fully-connected، لایه Softmax، و لایه طبقه‌بندی استفاده گردید. معماری کلی شبکه پیشنهادی به شبکه‌های معروفی مانند AlexNet شباهت دارد. اما جزئیات معماری شبکه با آزمون و خطا بر روی داده EMG به دست آمدند. پیش از هر کانولوشن، داده‌های ورودی Zero-pad شدند. تعداد فیلترها در پنج لایه کانولوشن به ترتیب ۱۶، ۳۲، ۶۴، ۶۴ و ۱۶، و تمام فیلترها ۳×۳ هستند. پولینگ در نواحی ۲×۲ با stride=2 اعمال شد. همچنین الگوریتم SGD برای آموزش شبکه استفاده شد و نرخ یادگیری اولیه ۰/۰۰۱ در نظر گرفته شد. اندازه Minibatchها ۱۲۸ تعیین گردید و تعداد Epochها ۴۰ در نظر گرفته شد. شبکه بر روی یک Nvidia GTX 1050 GPU آموزش داده شد و آموزش برای هر سابجکت کمابیش ۱۰۰ ثانیه به طول انجامید.

برای روش بر پایه SVM، پنج ویژگی شامل مجموعه TD به علاوه فرکانس میانگین، از هر پنجره داده‌های EMG استخراج شد. اضافه کردن فرکانس میانگین به دلیل آن بود که مشاهده شد که سبب افزایش دقت طبقه‌بندی می‌شود. از کتابخانه libSVM متلب برای شبیه‌سازی SVM استفاده گردید. کرنل Radial-basis و $\gamma = 1/40$ و $\text{cost}=1$ در نظر گرفته شدند. این مقادیر بهینه با آزمون و خطا به دست آمدند. برای به دست آوردن دقت طبقه‌بندی در دو روش CNN و SVM از 4-fold cross-validation استفاده شد. برای مقایسه دقت طبقه‌بندی بین دو روش CNN و SVM، از یک آزمون نمونه‌های جفت t-test استفاده شد.

یافته‌ها

شکل ۳ (الف) دقت طبقه‌بندی دو روش SVM و CNN را به همراه انحراف معیار بین ۱۰ کاربر، نشان می‌دهد. شکل ۳ (ب) دقت طبقه‌بندی برای هر کاربر را برای هر روش به تصویر می‌کشد. جدول ۱، یافته‌های آزمون نمونه‌های جفت t-test را نشان می‌دهد. براساس این یافته‌ها، تفاوت معناداری بین دقت طبقه‌بندی روش SVM و CNN به دست نیامد ($P=0/19$).

STOP

Extension & Pronation



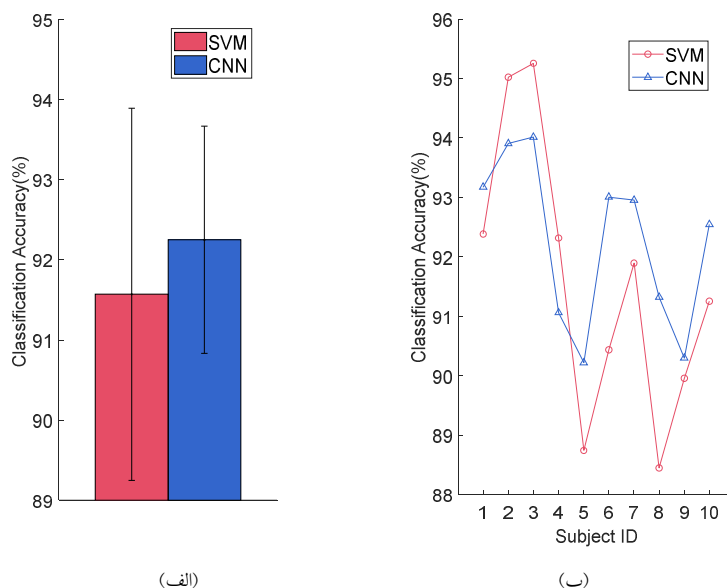
شکل ۲: یک اسکرین‌شات از اینترنت آموزش، در ترایال حرکت ترکیبی اکستنشن-پرونیشن

پیشرفت (Progress-bar) بود، به طوری که به عنوان نمونه اگر این خط به اندازه ۳۰٪ پر بود، کاربر باید با شدت ۳۰٪ نیروی کامل، آن انقباض را انجام دهد. هر ترایال شامل چهار تکرار از این سیکل بود: ۳ ثانیه حالت استراحت، ۳ ثانیه حرکت به سمت انقباض کامل، ۳ ثانیه نگه داشتن انقباض کامل، و ۳ ثانیه حرکت برگشتی به سمت حالت استراحت.

در هر ترایال، سیگنال EMG و هدف دیداری متناظر آن ضبط شدند. فرکانس نمونه‌برداری ۱/۲ کیلوهرتز برای ضبط EMG به کار گرفته شد. داده‌های EMG با یک فیلتر میانگذر باترورث درجه ۸ بین ۵-۵۰۰ هرتز فیلتر شدند. همچنین فرکانس ۵۰ هرتز توسط یک ناچ فیلتر باترورث مرتبه ۸ حذف گردید. ضبط و پردازش داده‌ها با نرم‌افزار متلب 2017Rb انجام شد.

دو روش یادگیری ماشین برای تشخیص اراده حرکتی از روی EMG به کار گرفته شد: ۱- روش بر پایه SVM (که عملکرد بالای آن در مطالعات پیشین^{۱۳} نشان داده شده است)، ۲- روش پیشنهادی بر پایه CNN. ابتدا برای ۸ ترایال اکتیو، قسمت‌هایی که هدف دیداری (و بنابراین انقباض) بالاتر از ۵۰٪ نیروی کامل بودند، نگه داشته شدند و بقیه قسمت‌ها حذف شدند. ۹ کلاس در نظر گرفته شد که شامل ۸ حرکت بیان شده در بالا و حالت استراحت بودند. سپس، داده‌های EMG به پنجره‌های زمانی با طول ۱۶۷ میلی‌ثانیه (۲۰۰ نمونه زمانی) با پرش ۴۰ میلی‌ثانیه تقسیم شدند.

برای روش بر پایه CNN، داده‌های EMG به صورت ماتریس‌های ۸×۲۰۰ مرتب شدند (۸: تعداد کانال‌ها، ۲۰۰: تعداد نمونه‌ها در هر پنجره زمانی). این ماتریس‌ها و برچسب نوع انقباض متناظر آن‌ها برای آموزش دادن شبکه CNN استفاده شدند. CNN پیشنهادی ۲۲ لایه



شکل ۳: دقت طبقه‌بندی روش‌های SVM و CNN: (الف) میانگین و انحراف معیار در بین تمام کاربرها (ب) برای هر کاربر

جدول ۱: یافته‌های آزمون آماری نمونه‌های جفت *t-test* برای مقایسه دو روش SVM و CNN: مقادیر *t*، *P*، درجه آزادی (*df*)، و انحراف معیار (*sd*) تفاضل دو گروه نشان داده شده‌اند

| | SVM | CNN | <i>p</i> | <i>t</i> | <i>df</i> | <i>sd</i> |
|-----------------------------|------------|------------|----------|----------|-----------|-----------|
| Classification Accuracy (%) | 91.6 ± 2.3 | 92.3 ± 1.4 | 0.19 | -1.4 | 9 | 0.02 |

بحث

نیروی حاصل از یک عضله به ویژگی‌های زمانی و فرکانسی سیگنال EMG ارتباط دارد. به همین دلیل، ویژگی‌های مجموعه TD هم شامل اطلاعات زمانی و فرکانسی سیگنال می‌باشد. عملکرد بالای CNN پیشنهادی نشان می‌دهد که برای تخمین اراده حرکتی از روی EMG، مدل CNN توانایی استخراج و یادگیری اطلاعات زمانی و فرکانسی لازم از داده‌های خام EMG را دارد. در مطالعات آینده، مقاومت CNN به دخالت‌های خارجی در سیگنال EMG مانند اثر جابه‌جایی الکترودها و تغییر حالت قرارگیری دست و ساعد بررسی خواهد شد. همچنین، با توجه به توانایی CNN در پردازش داده‌های با ابعاد بالا، ترکیب سیگنال‌های EMG و Force-myography (FMG) برای تشخیص اراده حرکتی مورد بررسی قرار خواهد گرفت.

این مطالعه، کارایی CNN برای تشخیص حرکات انفرادی و ترکیبی مچ از روی EMG را ارزیابی کرد. سیستم پیشنهادی می‌تواند در دست‌های مایوالکتریک متحرک به کار گرفته شود. براساس یافته‌ها، CNN به عملکردی مشابه SVM دست یافت که نشانگر کارایی بالای CNN پیشنهادی می‌باشد. مزیت روش پیشنهادی این است که نیازی به مهندسی ویژگی ندارد. از آنجاکه برای روش SVM از ویژگی‌های معروف و پرکاربرد EMG استفاده شد، یافته‌های این مطالعه حاکی از آن است که CNN توانایی یادگیری ویژگی‌های مفید از روی داده‌های خام EMG، برای طبقه‌بندی را دارا می‌باشد. از نقطه نظر فیزیولوژی،

References

1. Farina D, Jiang N, Rehbaum H, Holobar A, Graimann B, Dietl H, et al. The extraction of neural information from the surface EMG for the control of upper-limb prostheses: emerging avenues and challenges. *IEEE Trans Neural Syst Rehabil Eng* 2014;22(4):797-809.
2. Bishop CM. *Pattern Recognition and Machine Learning*. New York, NY: Springer-Verlag; 2016.
3. Englehart K, Hudgins B. A robust, real-time control scheme for multifunction myoelectric control. *IEEE Trans Biomed Eng* 2003;50(7):848-54.
4. Hudgins B, Parker P, Scott RN. A new strategy for multifunction myoelectric control. *IEEE Trans Biomed Eng* 1993;40:82-94.
5. Phinyomark A, Phukpattaranont P, Limsakul C. Feature reduction and selection for EMG signal classification. *Exp Syst Appl* 2012;39(8):7420-31.
6. Scheme E, Englehart K. Electromyogram pattern recognition for control of powered upper-limb prostheses: state of the art and challenges for clinical use. *J Rehabil Res Dev* 2011;48(6):643-59.
7. Phinyomark A, Khushaba RN, Ibáñez-Marcelo E, Patania A, Scheme E, Petri G. Navigating features: a topologically informed chart of electromyographic features space. *J R Soc Interface* 2017;14(137). pii: 20170734.
8. Goodfellow I, Bengio Y, Courville A, Bengio Y. *Deep Learning*. Cambridge: MIT press; 2016.
9. Atzori M, Cognolato M, Müller H. Deep learning with convolutional neural networks applied to electromyography data: a resource for the classification of movements for prosthetic hands. *Front Neurobot* 2016;10:9.
10. Geng W, Du Y, Jin W, Wei W, Hu Y, Li J. Gesture recognition by instantaneous surface EMG images. *Sci Rep* 2016;6:36571.
11. Xia P, Hu J, Peng Y. EMG-Based Estimation of Limb Movement Using Deep Learning With Recurrent Convolutional Neural Networks. *Artif Organs* 2018;42(5):E67-E77.
12. Cote-Allard U, Fall CL, Drouin A, Campeau-Lecours A, Gosselin C, Glette K, et al. Deep learning for electromyographic hand gesture signal classification using transfer learning. *IEEE Trans Neural Syst Rehabil Eng* 2019;27(4):760-71.
13. Ameri A, Kamavuako EN, Scheme EJ, Englehart KB, Parker PA. Support vector regression for improved real-time, simultaneous myoelectric control. *IEEE Trans Neural Syst Rehabil Eng* 2014;22(6):1198-209.

EMG-based wrist gesture recognition using a convolutional neural network

Ali Ameri Ph.D.*

Department of Biomedical
Engineering, School of Medicine,
Shahid Beheshti University of
Medical Sciences, Tehran, Iran.

* Corresponding author: School of
Medicine, Shahid Beheshti University of
Medical Sciences, Velenjak, Tehran,
Iran.
Tel: +98 21 22439941
E-mail: ameri@sbmu.ac.ir

Abstract

Received: 07 Aug. 2019 Revised: 14 Aug. 2019 Accepted: 12 Oct. 2019 Available online: 22 Oct. 2019

Background: Deep learning has revolutionized artificial intelligence and has transformed many fields. It allows processing high-dimensional data (such as signals or images) without the need for feature engineering. The aim of this research is to develop a deep learning-based system to decode motor intent from electromyogram (EMG) signals.

Methods: A myoelectric system based on convolutional neural networks (CNN) is proposed, as an alternative to conventional classification methods that depend on feature engineering. The proposed model was validated with 10 able-bodied subjects during single and combined wrist motions. Eight EMG channels were recorded using eight pairs of surface electrodes attached around the subject's dominant forearm. The raw EMG data from windows of 167ms (200 samples) in 8 channels were arranged as 200×8 matrices. For each subject, a CNN was trained using the EMG matrices as the input and the corresponding motion classes as the target. The resulting model was tested using a 4-fold cross-validation. The performance of the proposed approach was compared to that of a standard SVM-based model that used a set of time-domain (TD) features including mean absolute value, zero crossings, slope sign changes, waveform length, and mean frequency.

Results: In spite of the proven performance and popularity of the TD features, no significant difference ($P=0.19$) was found between the classification accuracies of the two methods. The advantage of the proposed model is that it does not need manual extraction of features, as the CNN can automatically learn and extract required representations from the EMG data.

Conclusion: These results indicate the capacity of CNNs to learn and extract rich and complex information from biological signals. Because both amplitude and frequency of EMG increases with increasing muscle force, both temporal and spectral characteristics of EMG are needed for efficient estimation of motor intent. The TD set, also includes these types of features. The high performance of the CNN model shows its capability to learn temporal and spectral representations from raw EMG data.

Keywords: classification, convolutional neural network, deep learning, electromyogram.