

Deep Learning-Based Classification Using Cumulants and Bispectrum of EMG Signals

E. Orosco, J. Gaia, J. Gimenez, and C. Soria

Abstract—Surface electromyographic signals (EMG) historically have been used to classify tasks in basis of a feature extraction scheme and low complexity classifiers. Deep networks, as Multilayer Perceptron and Convolutional Neural Network (MLP and CNN, respectively), avoid the traditional, complex and heuristic (handcrafted) process of feature extraction. Today, it is possible to face the computational cost that these automatic techniques require due to the technology advancement. This allowed deep learning techniques to be quickly generalized to countless applications. This paper proposes to use the third order cumulants and their 2D Fourier transform (Bispectrum) to directly feed CNN and MLP deep learning networks. The classifier is not user-dependent (same classifier for all users) and obtains better results than the classical scheme according to several metrics.

Index Terms—EMG, Cumulants, Bispectrum, CNN, MLP.

I. INTRODUCCIÓN

Las regulaciones de movimiento en el sistema motor pueden representarse como sinergias musculares, las cuales especifican un patrón de activación muscular, donde se supone que el nivel absoluto de activación de cada sinergia muscular está modulado por señales de comando neuronal [1]. Esta manifestación se conoce como señal electromiográfica (EMG) y puede adquirirse desde la superficie de la piel de forma no invasiva [2], [3]. La fisiología del cuerpo humano usa tantas activaciones neuronales como sea necesario, generando una red neuronal densa que permite la sinergia muscular. La evolución de las técnicas de aprendizaje profundo (Deep Learning [4]), permite estudiar problemas tan complejos como las sinergias musculares.

Históricamente, hay dos áreas principales relacionadas con el procesamiento de señales EMG: la extracción de características y la clasificación. Existe un gran número de procedimientos clásicos que permiten extraer características representativas de señales EMG [5], [6], [7], [8], [9]. Entre las características utilizadas para el procesamiento de señales EMG se encuentran las basadas en estadísticas de alto orden (sigla en inglés, HOS), donde se destacan los integradores sobre segmentos definidos en la matriz del bispectro [10], la media del módulo de las entradas del bispectro [11], y medidas de posición calculadas con base en módulo o el componente real de la matriz bispectro [12]. Los HOS han llegado a utilizarse incluso para modelar señales EMG [13].

Las técnicas basadas en HOS no son características, sino que son una representación alternativa a la secuencia temporal

de las señales EMG. El bispectro es la transformada de Fourier 2D de los cumulantes de tercer orden. De forma análoga que la autocorrelación y el espectro de potencia, los cumulantes de tercer orden y el bispectro brindan información de las señales EMG en dos dominios distintos. Generalmente, a las señales EMG se les adjudica una distribución estadística Gaussiana, que presenta valores nulos en la secuencia de cumulantes de orden superior al segundo. La ventaja de utilizar HOS es que describen aspectos de las señales no tan evidentes para las estadísticas de segundo orden [12], [14], [15]. Se han relacionado las estadísticas de alto orden con el estudio de los procesos estocásticos no Gaussianos. Si el proceso estocástico es Gaussiano, sus funciones de densidad de probabilidad marginales y conjuntas son simétricas, entonces el cumulante enésimo es nulo. Los multiespectros de orden mayor a dos son cero en presencia de distribuciones Gaussianas (marginales y conjuntas), por lo que multiespectros no nulos proveen información adicional respecto a la desviación no Gaussiana de la distribución de probabilidad en los procesos estocásticos [16], [17], [18], [19].

En cuanto a la descomposición e interpretación de sinergias musculares, se ha utilizado análisis factorial (FA) con el análisis de componentes independientes [20], [21], [22], [23], [24] y la factorización de matriz no negativa [25]. En los trabajos [12], [26], [14], [15], el equipo de investigación abordó el problema de la clasificación de características basadas en HOS con el fin de realizar control mioeléctrico. En [12] se formulan varios vectores de características basados en el bispectro para luego clasificar con una exactitud comprendida entre 90% y 97%. En [14] se realizó una comparación entre una característica basada en bispectro y características basada en cumulantes de tercer orden. En [15] la propiedad de sinergias entre grupos musculares justificó que los cumulantes cruzados de tercer orden fueran utilizados como características para clasificar las actividades planteadas. Para la etapa de clasificación, en todos los trabajos mencionados, se han implementado diversas técnicas como redes neuronales, máquinas de soporte vectorial, e incluso clasificadores lineales.

En contraposición con las técnicas clásicas de procesamiento, el aprendizaje profundo permite generar procesos novedosos de clasificación que no requieren de una etapa previa de extracción de características, la cual suele ser una tarea artesanal. Existen pocos trabajos que incursionan en clasificar tareas con señales EMG sin extraer características, de los cuales la mayoría se publicaron en los últimos dos años [27]. El trabajo [28] propone un sistema de control mioeléctrico basado en redes neuronales convolucionales (CNN), el cual se compara con una máquina de soporte vectorial estándar con características en el dominio temporal. En [29] se desarrolla

Dr. E.C. Orosco (eorosco@inaut.unsj.edu.ar), Ing. J. Gaia (jere-gaia2014@gmail.com), Dr. J. Gimenez (jgimenez@inaut.unsj.edu.ar), Dr. C. Soria. (eorosco@inaut.unsj.edu.ar). Instituto de Automática de la Facultad Ingeniería, UNSJ-CONICET, San Juan, Argentina.

una tesis que compara el rendimiento entre algoritmos de aprendizaje profundo (una CNN, una RNN, y una combinación secuencial de ambas) para el control mioeléctrico fuera de línea. En [30] se propone el uso de señales EMG sin procesar como entradas directas a una CNN. En [31] se clasifican movimientos en normales y agresivos utilizando dos representaciones tiempo-frecuencia (el espectrograma y el escalograma) de señales EMG como entradas de tres CNNs. Todos los casos concluyen en que no se encuentran diferencias de rendimiento entre utilizar clasificadores clásicos con extracción de características, o usar directamente técnicas de aprendizaje profundo con datos sin procesar. Sin embargo, el diseño correcto del vector de características suele ser más crítico que la elección del método de clasificación [32]. Este paradigma está siendo refutado con el advenimiento del aprendizaje profundo.

En este trabajo se propone emplear aprendizaje profundo basado en representaciones HOS de las señales EMG, de lo que no existe antecedentes en la literatura. El objetivo principal es reemplazar la extracción de características heurística por un método automático (aprendizaje profundo), y realizar un estudio comparativo entre los resultados aquí obtenidos y las técnicas basadas en características desarrolladas anteriormente en [12], [14], [15], investigando cuán significativo puede ser usar la transformada Fourier 2D o no para la tarea de clasificación de movimientos musculares.

El trabajo se organiza del siguiente modo: En la Sección II se dan detalles sobre la obtención de la base de datos, sobre el preprocesamiento aplicado, y sobre la transformación de los datos propuesta para realizar el entrenamiento. Además, se presenta una vista general sobre las arquitecturas de redes neuronales utilizadas para clasificar los movimientos. En la Sección III se realizan diversas experimentaciones combinando como entradas las estimaciones de los cumulantes de tercer orden o la estimación del biespectro, con las arquitecturas de redes neuronales densas perceptrón multicapas (MLP) y convolucionales (CNN). Además, se muestran resultados comparativos respecto a las tasas de clasificación obtenidas en trabajos previos basados en características. Finalmente, la Sección IV contiene las conclusiones del trabajo.

II. MATERIALES Y MÉTODOS

En esta sección se procede a explicar la ejecución de los protocolos para la obtención de la base de datos originada en [12]. Posteriormente se resume la teoría de la representación de las señales EMG en cumulantes de tercer orden y en biespectro, y se da una breve explicación de las MLP y CNN.

A. Protocolo EMG

El protocolo está planteado para registrar señales EMG de los músculos bíceps y tríceps braquial, pronador y braquiorradial (ver Fig. 1). Los movimientos musculares estudiados para las tareas de clasificación son flexión, extensión, pronación, supinación, agarre y posición inactiva (sin contracciones). Seis sujetos sanos (3 hombres y 3 mujeres, entre las edades de 23 y 34 años y rango de peso 53.5-78.9kg) son guiados para

ejecutar determinados movimientos. Uno de los sujetos tiene una focomelia unilateral debajo de su codo. Estos sujetos, usuarios finales del sistema desarrollado, firmaron y aprobaron un formulario de *consentimiento informado* de acuerdo con el protocolo. La base de datos se recopiló para el trabajo [12] (el movimiento de agarre no se consideró en ese trabajo), y luego, se usó en [14] y [15]. No se requirió de la aprobación de un comité de ética por el alcance del trabajo. El registro de datos se realizó durante 4 días para prevenir fatigas musculares. Cada día, los sujetos ejecutaron una prueba experimental que consta de cinco series de movimientos secuenciales compuestos de 3s de una contracción específica intercalada por 2 a 3s de inactividad, totalizando 120s de duración. Las señales EMG se digitalizaron a una frecuencia de 1kHz (1ms de tiempo de muestreo). Se adquirieron $120s \times 1.000$ muestras por segundo $\times 4$ canales $\times 6$ sujetos $\times 4$ días. Siendo 2.880.000 muestras multi variables (considerando los 4 canales), totalizando 11.520.000 datos.

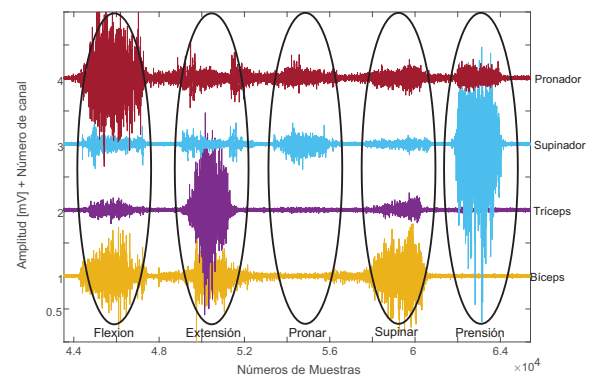


Fig. 1. Ejemplo de señales EMG para un sujeto sano. Registro proveniente de los músculos bíceps y tríceps braquial, pronador y braquiorradial. Los movimientos musculares a clasificar son flexión, extensión, pronación (pronar), supinación (supinar), agarre (o prensión) y posición inactiva (sin contracciones).

B. Segmentación de Datos

Cada señal EMG se particiona en segmentos de 256 muestras (256ms) con solapamientos entre segmentos contiguos de 128 muestras, de manera que el clasificador cuente con información temporal y contextual de los datos que recibe. El tamaño del segmento presenta una relación de compromiso tal que, mientras menor sea, más rápida es la respuesta y, por otro lado, mientras más grande, menor es el sesgo y varianza de los parámetros estimados. Nótese que, si el tamaño del segmento es muy grande, aumenta la probabilidad de abarcar más de un movimiento. En la Fig. 2 se diagrama el proceso típico de segmentación sobre una señal EMG. De este proceso se desprende que el número resultante de muestras segmentadas es de aproximadamente 2196 para cada una de las cuatro señales EMG.

C. Estadísticas de Alto Orden

Los procesos estocásticos pueden ser caracterizados estadísticamente a través de momentos y cumulantes de bajo

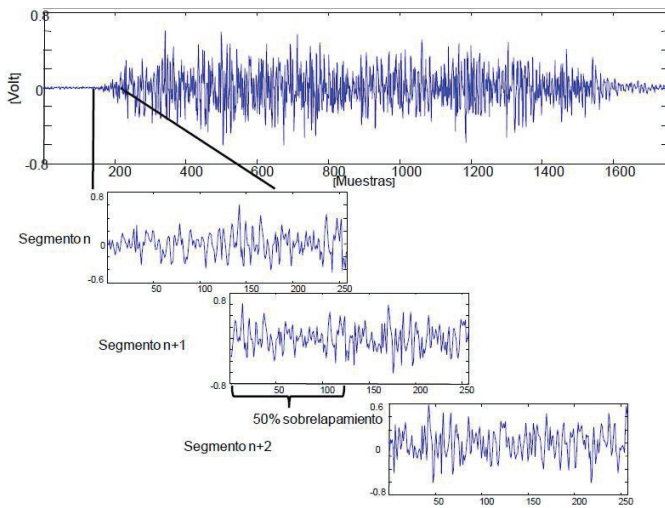


Fig. 2. Proceso típico de segmentación sobre una señales EMG.

y de alto orden en el dominio temporal y frecuencial. La media, varianza y covarianza son de bajo orden, mientras que la asimetría y curtosis son estadísticos de alto orden. En el caso univariado la función de momento de n -ésimo orden se define como:

$$m_n(\tau_1, \tau_2, \dots, \tau_{n-1}) \triangleq E\{X(k)X(k+\tau_1)\dots X(k+\tau_{n-1})\}, \quad (1)$$

donde cada τ_i es un desplazamiento temporal y $E\{\}$ es el operador esperanza estadística. Casos particulares de (1) son la autocorrelación o momento de segundo orden $m_2(\tau)$, y el momento de tercer orden $m_3(\tau_1, \tau_2)$. Para $n = 1$ y $n = 2$, se tienen la media y varianza del proceso $X(k)$, y para $n = 3$, cuando el proceso es de media cero ($E\{X(k)\} = 0 \forall k$), representa la función del cumulante de tercer orden $c_3(\tau_1, \tau_2)$, esto es:

$$c_3(\tau_1, \tau_2) = m_3(\tau_1, \tau_2) = E\{X(k)X(k+\tau_1)X(k+\tau_2)\}. \quad (2)$$

Este trabajo se centra en el cumulante de tercer orden (ver detalles en [12], [14], [15], [16], [33]).

Se define el biespectro como la transformada de Fourier 2D del cumulante de tercer orden:

$$B_3(\omega_1, \omega_2) = \sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} c_3(\tau_1, \tau_2) \exp\{-j(\tau_1\omega_1 + \tau_2\omega_2)\}, \quad (3)$$

donde $|\omega_1| \leq \pi$; $|\omega_2| \leq \pi$; $|\omega_1 + \omega_2| \leq \pi$. Las propiedades de simetrías son semejantes a las de bajo orden (ver detalles en [16], [33]).

D. Estimación de Cumulantes de Tercer Orden

En [14] se estudiaron tres estimadores del cumulante de tercer orden (2), basados en medias, medianas, y medias truncadas, respectivamente. Aquí se utiliza la mediana por sus propiedades de menor varianza y robustez a valor atípicos, esto es:

$$\hat{c}_3(m, n) = \text{mediana}(\{X(l)X(l+n)X(l+m)/l = 1 : N-m\}) \quad (4)$$

para cada m y n de la región triangular $0 \leq n \leq m \leq N-1$, siendo N el tamaño de la muestra.

Para concretar estas estimaciones y con el fin de reducir variabilidades, cada segmento diagramado en la Fig. 2 se particiona en 4 subsegmentos, de los cuales se obtienen cuatro estimaciones de $\hat{c}_3(m, n)$ con $N = 64$. Al promediar estas cuatro estimaciones, se obtiene una matriz de 64×64 por cada segmento de la señal EMG. Si estas estimaciones se hacen simultáneamente en los cuatro canales (estimaciones basadas en los músculos Bíceps, Tríceps, Supinador y Pronador), y se concatenan las estimaciones correspondientes al mismo tiempo, se puede conformar una matriz de 128×128 por cada segmento de tiempo. Para más detalles ver la Fig. 3, donde se muestra la estimación del cumulante de tercer orden para cinco movimientos musculares (flexión, extensión, pronación, supinación y agarre). Estas matrices de 128×128 son las utilizadas para las experimentaciones en general, mientras que las matrices sin concatenar de 64×64 solo se utilizan en la Sección III-D, donde se realiza un análisis mono-canal.

E. Estimación de Biespectro

El problema de estimar el biespectro a partir de una sola realización de un proceso estocástico se ha estudiado ampliamente en la literatura. Básicamente, se han utilizado dos enfoques: el enfoque paramétrico (que se basa en modelos paramétricos lineales como los modelos autorregresivos, de media móvil y ARMA); y el enfoque convencional o “tipo de Fourier” [16]. La precisión de la estimación del biespectro depende principalmente de la precisión de la estimación del cumulante de tercer orden. La estimación del biespectro se reduce a aplicar propiedades de simetría de los cumulantes y propiedades de la transformada rápida de Fourier 2D. Como en el caso de la estimación del espectro de potencia convencional, se utiliza una función de ventana bidimensional $w(m, n)$ para encontrar estimaciones suaves del biespectro, esto es:

$$B_3(\omega_1, \omega_2) = \sum_{m=-L}^L \sum_{n=-L}^L \hat{c}_3(m, n) w(m, n) \exp\{-j(\omega_1 m + \omega_2 n)\}, \quad (5)$$

De igual forma que con la estimación del cumulante de tercer orden, el resultado de la estimación del biespectro está compuesto por cuatro matrices de 64×64 , que se concatenan de igual forma que $\hat{c}_3(m, n)$ para conformar una matriz de 128×128 . En la Fig. 3 se muestra el módulo de las estimaciones para cinco movimientos musculares (flexión, extensión, pronación, supinación y agarre), las cuales se utilizarán para entrenar las redes MLP y CNN con el fin de clasificar movimientos musculares.

F. Arquitecturas de Clasificadores de Aprendizaje Profundo

Existen diversas arquitecturas y paradigmas dentro de los procesos de aprendizaje profundo [4]. En el presente trabajo se considerarán dos de ellos: las MLP y las CNN. Ambas redes se implementarán con el fin de aprender de las imágenes producidas por las estimaciones de los cumulantes y del biespectro de las señales EMG, las cuales no se consideran características

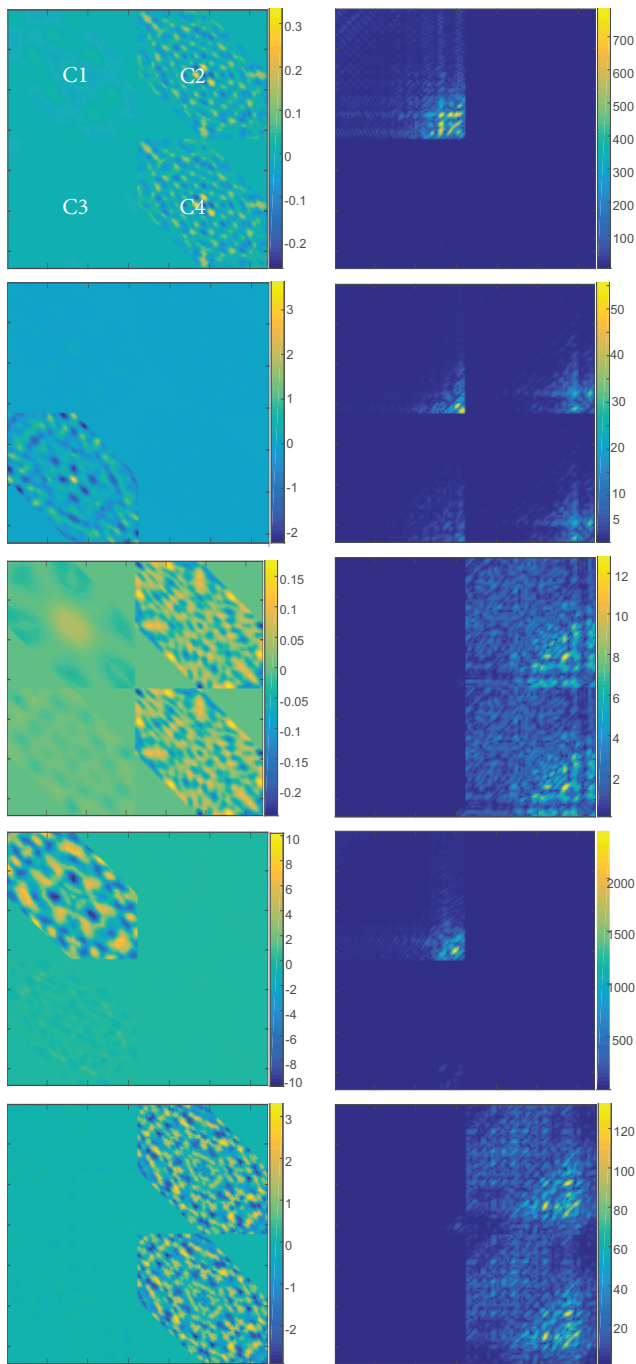


Fig. 3. Estimación del cumulante de tercer orden (columna izquierda) y del bispectro (columna derecha), basada en segmentos de señales EMG del Bíceps (C1), Tríceps (C2), Supinador (C3) y Pronador (C4), concatenadas en un arreglo matricial cuadrado. De arriba a abajo se muestran las acciones de flexión, extensión, pronación, supinación, y agarre.

sino transformaciones o representaciones transparentes de las señales.

1) *Red neuronal densa perceptrón multicapa (MLP)*: La arquitectura MLP difiere de la de una red neuronal artificial tradicional (ANN) en que posee mayor cantidad de capas y neuronas por capa, las cuales se encuentran totalmente conectadas. Esta estructura le provee a las MLP (Fig. 4), mayor robustez ante perturbaciones en los datos de entrada, y a la

vez, a medida que se incorporan más capas, las neuronas de las capas lejanas respecto de la entrada se activan de acuerdo a características cada vez más complejas. Generalmente, estas características no son triviales e intuitivas como para ser propuestas de antemano por procedimientos heurísticos. Por estos motivos es que se dice que las MLP aprenden automáticamente las características óptimas para el problema particular tratado. El proceso de entrenamiento de las MLP es el mismo que el de las ANN, y por ende, requieren de un costo computacional considerablemente mayor. El reciente aumento en las aplicaciones de las MLP se debe en parte a los avances en la tecnología de GPU que han reducido sustancialmente el tiempo de entrenamiento del modelo. Una dificultad limitante de las MLP es que el gradiente de la función de costo que se optimiza durante el entrenamiento es insignificante en las entradas correspondientes a los pesos de las capas lejanas, y por ende, el proceso de ajuste es muy lento. Además, la dimensionalidad de la red debe controlarse para no incurrir en problemas de sobreajuste. Para atenuar estas dificultades, las MLP incorporan un proceso de eliminación aleatoria de activaciones (dropout) que le permite no quedar atrapado en mínimos locales. Otra estrategia implementada para tal fin es regular la tasa de aprendizaje (learning rate).

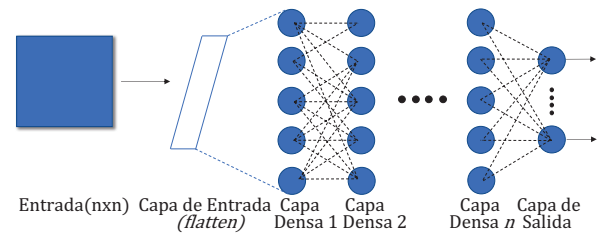


Fig. 4. Arquitectura de redes MLP.

La estructura elegida con criterios de maximizar la precisión y minimizar la estructura cuenta con una capa de entrada (Flaten), dos capas internas densas (Dense), un proceso de eliminación aleatoria de activaciones (dropout), y una capa de salida densa (Dense). El tamaño de las capas internas son iguales y se investigan como criterio de elección de una MLP de máxima precisión.

2) *Redes Neuronales Convolucionales (CNN)*: Las entradas a una MLP son vectores de datos que se asumen descorrelacionados, lo cual produce una considerable pérdida de información contextual. Las CNN se diseñaron con el fin de aprovechar las características espaciales propias de las imágenes, las cuales son ignoradas por las demás arquitecturas. Estas redes involucran tres tipos principales de capas: las capas convolucionales, las capas de agrupación (pooling) y las capas densas (ver Fig. 5). Cada capa contiene diversas características que pueden ajustarse de acuerdo a los requerimientos del problema en cuestión. Las capas convolucionales y las de agrupación se alternan con el fin de resaltar determinadas características, y a la vez, controlar la cantidad de datos que ingresarán posteriormente a la capa densa encargada de realizar la tarea de síntesis y clasificación. Esta fase de extracción de características se asemeja al proceso estimulante en las células de la corteza visual. En cada capa convolucional

se le aplican diversos filtros convolucionales a la imagen de entrada. Las salidas de esta capa son tantas imágenes como filtros se apliquen por cada imagen de entrada a la capa. El tamaño de estas imágenes de salida son del orden del tamaño de las imágenes de entrada, y depende del tamaño del filtro (generalmente de 3×3 o 5×5) y del paso que se realice durante la convolución (stride length). Luego de este proceso de filtrado los datos se multiplican considerablemente, y por ello, las capas convolucionales se alternan con capas de agrupamiento, donde cada imagen es particionada regularmente en subimágenes (generalmente de 2×2 o 3×3), para luego reducir cada una de estas a un sólo dato (el promedio o el máximo de las entradas de la submatriz). Los filtros son entrenados automáticamente con el fin de producir un mapa de características óptimo, los cuales se utilizan como entradas de la capa densa final encargada de realizar la clasificación. La Tabla I expone una lista detallada de las configuraciones establecidas para CNN.

TABLA I
ESTRUCTURA Y TAMAÑO DE LAS CNN

Capa	Tamaño	Parámetros
Convocional (Conv2D)	(128, 128, 32)	320
Activación	(128, 128, 32)	0
Convocional (Conv2D)	(126, 126, 32)	9248
Activación	(126, 126, 32)	0
Reducción (MaxPooling)	(63, 63, 32)	0
Descarte (Dropout)	(63, 63, 32)	0
Lineal (Flatten)	127008	0
Interna (Dense)	512	65028608
Activación	512	0
Descarte (Dropout)	512	0
Salida (Dense)	6	3078
Total	Parámetros	65041254

III. RESULTADOS

Esta sección describe los experimentos llevados a cabo para evaluar las estructuras de aprendizaje profundo MLP y CNN, alimentadas de cumulantes de tercer orden y del biespectro de acuerdo a las Sección II-A y II-B. Se particiona aleatoriamente las muestras dejando un 70% para entrenar y un 30% para validar. Los ensayos experimentales son realizados en búsqueda de maximizar la precisión y minimizar la estructura. Las Secciones II-F1 y II-F2 exponen los detalles de las configuraciones establecidas para MLP y CNN, respectivamente.

El rendimiento de cada red se estimó utilizando cuatro tipos de mediciones de evaluación: sensibilidad (6), precisión (7), f1-score (8), y exactitud (9), las cuales se calculan por clase y se basan en las cantidades de verdaderos positivos (VP), verdaderos negativos (VN), falsos positivos (FP) y falsos negativos (FN).

$$S = VP/(VP + FN), \quad (6)$$

$$P = VP/(VP + FP), \quad (7)$$

$$F = 2VP/(2VP + FP + FN), \quad (8)$$

$$E = (VP + VN)/(VP + VN + FP + FN), \quad (9)$$

Las opciones de entrenamiento fueron elegidas con el fin de evitar el sobreentrenamiento. Se utiliza el algoritmo *Adam* de optimización de velocidad de aprendizaje adaptativo diseñado específicamente para entrenar redes neuronales profundas, con los parámetros recomendados en [34], esto es, $\alpha = 0.001$, $\beta_1 = 0.9$, $\beta_2 = 0.999$ y $\epsilon_1 = 10^{-8}$. La función de pérdida (o función de puntaje de optimización) es elegida por entropía cruzada (sparse categorical crossentropy) con etiquetas categóricas del 0 al 5. El número de iteraciones se estableció experimentalmente según la convergencia de la precisión, quedando en 50 para MLP y en 20 para CNN, con tiempos de procesamiento aproximados de 700.39s y 1241.05s, respectivamente. Para ejecutar los proyectos de aprendizaje automático se utiliza Jupyter Notebook en Google Colaboratory o Colab, que proporciona una plataforma lo suficientemente potente. Se usa de forma gratuita un GPU NVIDIA Tesla T4, cargando los datos en Google Drive y se abre un cuaderno en Colaboratory.

A. Determinación del Tamaño de Capas Internas de MLP

En esta sección se determinan el tamaño de capas internas de MLP con cumulantes de tercer orden y con biespectro. Como se detalla en la Sección II-F1, se utilizan dos capas internas de igual tamaño, el cual se varía entre 64, 128, 256, 512 y 1024 para investigar su efecto en las mediciones de evaluación. En la Fig. 6 se evidencia un crecimiento sostenido en todas las mediciones (S, P, F, E) al aumentar el tamaño desde 64 hasta 512, mientras que no se evidencian cambios significativos al aumentar el tamaño de 512 a 1024. Por éste motivo es que, en las experimentaciones que se presentan a continuación, se escoje el tamaño de 512 para las capas internas de las redes MLP.

B. Resultados Obtenidos con Cumulantes de Tercer Orden

En este apartado las redes MLP y CNN son evaluadas con entrada de los cumulantes de tercer orden. La estructura MLP de la Sección II-F1 se usa con un tamaño de 512 y la red densa CNN según se expone en la Sección II-F2. En la Tabla II se comparan los resultados de las dos redes a través de los promedios de las mediciones de evaluación obtenidas para las distintas clases de movimientos. El menor índice está dado en la sensibilidad (S) para MLP (0.947179), mientras que el mayor lo alcanza en precisión (P) la red CNN con 0.991708. Los resultados en promedio son similares, aunque se evidencian diferencias cuando se discriminan por movimiento (comparar Tablas III y IV). El menor valor de sensibilidad (S) es de 0.889467 (MLP-Extensión), contra 0.975293 (CNN-Extensión). El valor más alto de sensibilidad es CNN-Inactivo de 0.996439. Esta diferencia de las CNN respecto a las MLP se fundamenta en el hecho de que las entradas son imágenes, las cuales son tratadas de mejor manera por las CNN que contemplan la información contextual y espacial de las mismas.

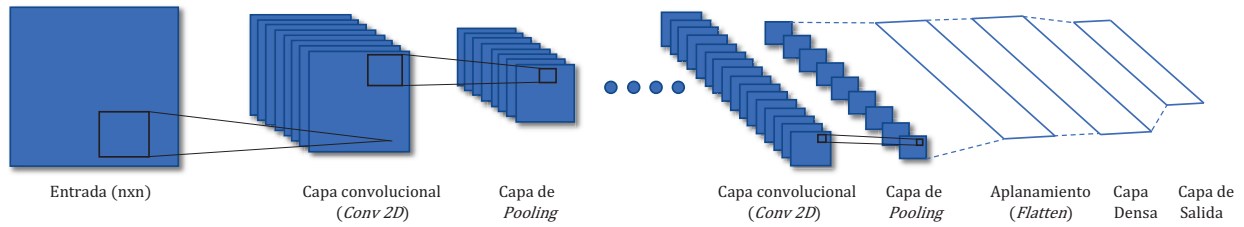


Fig. 5. Arquitectura de redes CNN.

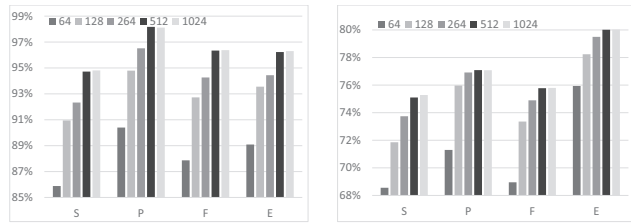


Fig. 6. Mediciones de evaluación S, P, F, E para diferentes tamaños de capas internas de MLP. Cumulantes de tercer orden (izquierda) y bispectro (derecha).

Tabla V se expone una comparación entre los resultados de las redes. Se observa una clara diferencia entre utilizar una CNN respecto a utilizar una MLP, las cuales se vieron seriamente perjudicadas al dejar de ser entrenadas con información en el dominio temporal. El menor índice está dado en la sensibilidad (S) para MLP (0.751012), mientras que CNN alcanzan un índice máximo de precisión de 0.989777. En las Tablas VI y VII se muestran los índices de clasificación discriminados por movimientos. Nótese que todos los movimientos son identificados con altas tasas de clasificación utilizando CNN, a diferencia de las MLP que muestran claras dificultades para trabajar con estos datos espectrales.

TABLA II

EVALUACIÓN CON CUMULANTES DE TERCER ORDEN, COMPARACIÓN MLP Y CNN

Evaluación	MLP	CNN
S	0.947179	0.980847
P	0.981731	0.991708
F	0.963436	0.986176
E	0.962267	0.985907

TABLA III

EVALUACIÓN DE LAS MLP, CON CUMULANTES DE TERCER ORDEN

Movimiento	P	S	F
Inactivo	0.926700	0.995014	0.959643
Pronar	0.982415	0.963218	0.972722
Supinar	0.995781	0.952894	0.973865
Flexión	0.995620	0.936813	0.965322
Extensión	0.989870	0.889467	0.936986
Prensión	0.999988	0.945668	0.972075

TABLA IV

EVALUACIÓN DE LAS CNN, CON CUMULANTES DE TERCER ORDEN

Movimiento	P	S	F
Inactivo	0.973895	0.996439	0.985038
Pronar	0.994206	0.986207	0.990190
Supinar	0.995924	0.986541	0.991210
Flexión	0.997167	0.967033	0.981869
Extensión	0.992063	0.975293	0.983607
Prensión	0.996992	0.973568	0.985141

TABLA V

PROMEDIOS DE EVALUACIÓN CON BIESPECTRO, COMPARACIÓN MLP Y CNN

Evaluación	MLP	CNN
S	0.751012	0.983386
P	0.770872	0.989777
F	0.757777	0.986530
E	0.800121	0.988180

TABLA VI

EVALUACIÓN DE LAS MLP, CON BIESPECTRO

Movimiento	P	S	F
Inactivo	0.880468	0.910256	0.895115
Pronar	0.777922	0.688506	0.730488
Supinar	0.956381	0.796770	0.869310
Flexión	0.690748	0.748626	0.718523
Extensión	0.763348	0.687906	0.723666
Prensión	0.556364	0.674009	0.609562

TABLA VII

EVALUACIÓN DE LAS CNN, CON BIESPECTRO

Movimiento	P	S	F
Inactivo	0.984880	0.997507	0.991154
Pronar	0.985160	0.991954	0.988545
Supinar	0.997279	0.986541	0.991881
Flexión	0.987535	0.979396	0.983448
Extensión	0.997361	0.983095	0.990177
Prensión	0.986446	0.961821	0.973978

C. Resultados Basados en el Biespectro

En esta sección se muestran los resultados de evaluación de las MLP y CNN utilizando como entrada el bispectro. En la

D. Resultados Basados en un Canal con CNN

Debido a que las CNN mostraron mejores resultados, en esta sección se presentan los resultados de las CNN utilizando cumulantes de tercer orden y biespectro basados solamente en un canal de la señal EMG a la vez. La estructura de la CNN es la especificada en la Tabla I, pero el tamaño de la matriz de entrada es de 64×64 , en vez de 128×128 , debido a que se usa sólo un canal (ver proceso de concatenación en Sección II-B). La Tabla VIII expone los resultados obtenidos, donde se observa una mayor dispersión y deterioro de los índices en general. Este resultado es de esperar debido a que la información compuesta que tienen todos los canales es mejor que la que contiene un solo canal. De igual manera, resulta loable que se tiene una precisión (P) promedio de 0.91 (con cuatro canales los índices de evaluación eran de 0.96-1.00), considerando que se cuenta con un sólo canal para clasificar seis clases. El deterioro también se ve reflejado en la sensibilidad (S), la cual es en promedio 0.83.

TABLA VIII
EVALUACIÓN DE LAS CNNs DISCRIMINADO POR CANAL

Tipo	Canal	S	P	F	E
Cumulante	Canal 1	0.82	0.95	0.87	0.88
Cumulante	Canal 2	0.69	0.84	0.73	0.78
Cumulante	Canal 3	0.81	0.90	0.85	0.87
Cumulante	Canal 4	0.82	0.94	0.87	0.87
Biespectro	Canal 1	0.79	0.89	0.83	0.85
Biespectro	Canal 2	0.87	0.89	0.88	0.89
Biespectro	Canal 3	0.89	0.95	0.92	0.92
Biespectro	Canal 4	0.89	0.92	0.90	0.91
Promedio		0.83	0.91	0.85	0.87

E. Comparación con Trabajos Previos y con VGG16

Como se menciona a lo largo del manuscrito, las técnicas estadísticas basadas en HOS (cumulantes y biespectro) ya han sido trabajadas de forma clásica por el grupo de investigación utilizando la misma base de datos. Se formularon dos métodos de reducción de características para la matriz de biespectro compleja [12]. En [14] se realizó una comparación entre una característica basada en biespectro y características basada en cumulantes de tercer orden. En ese trabajo, se utilizaron ANN y máquinas de vectores de soporte (SVM) para la detección de actividad muscular. Finalmente, en [15] se utilizaron los cumulantes cruzados de tercer orden como características para clasificar las actividades a través de ANNs particulares para cada usuario. En los trabajos anteriores se cuantificó la calidad de la clasificación utilizando solamente el índice de exactitud (E), y por ello, se elige como medida comparativa en este trabajo (ver Tabla IX). El planteo moderno de alimentar la CNN directamente con técnicas HOS es claramente superior al enfoque clásico de extracción de características basado en HOS.

Resulta natural utilizar otros modelos para comparar resultados, se estudiaron Inception V3, MobileNet y VGGNet, y se seleccionó por mejores resultados el modelo VGGNet en su versión VGG16 [35]. El método para la reutilización

del modelo pre-entrenado con los datos propios, se basa en el término conocido como transferencia de aprendizaje (siglas en inglés de transfer learning [36]). Esto consiste en adaptar la capa de entrada, agregar las capas más profundas (salida) y entrenar solo los parámetros nuevos conservando parcialmente los de la red original. VGG16 modificada posee un total de 16813638 parámetros de los cuales se entrenan 2098950 y se utilizaron 50 iteraciones (epoch). En la Tabla X se muestran los resultados comparativos entre cumulantes de tercer orden y biespectro evaluados a través de las métricas S, P, F y E. La exactitud (E) de cumulantes (0.808153) y de biespectro(0.8102749) resultan similares para este modelo pero inferiores a la red CNN propuesta por este trabajo.

TABLA IX
COMPARACIÓN DE TASAS DE EXACTITUD ENTRE ESQUEMAS CLÁSICOS Y POR APRENDIZAJE PROFUNDO

Transformada o Característica	Clasificador	E
Matriz Cumulante	MLP	0.96
Matriz Cumulante	CNN	0.98
Matriz Biespectro	MLP	0.80
Matriz Biespectro	CNN	0.98
Promedio uni-canal Cumulante-Biespectro	CNN	0.87
Máximo uni-canal Biespectro Canal 3	CNN	0.92
Característica basada en Biespectro*	ANN	0.92***
Característica basada en Cumulante*	ANN	0.90***
Característica basada en Biespectro**	Detector ANN	0.91***
Característica basada en Cumulante**	Detector ANN	0.90***
Característica basada en Biespectro**	Detector SVM	0.92***
Característica basada en Cumulante**	Detector SVM	0.91***

*Datos extraídos de Tabla 2 de [15]. **Datos extraídos de Tablas 2 y 3 de [14]. ***Promedio de clasificadores individuales.

TABLA X
EVALUACIÓN DE CUMULANTES DE TERCER ORDEN Y BIESPECTRO UTILIZANDO VGG16

Evaluación	Cumulantes	Biespectro
S	0.733692	0.737301
P	0.821608	0.828536
F	0.766614	0.775147
E	0.808153	0.810274

IV. CONCLUSION

En este trabajo se planteó de forma satisfactoria un esquema de clasificación por redes MLP y CNN entrenadas por transformadas de señales EMG. Se puede concluir que las redes CNN fueron superiores que las MLP en general, y similares en algún caso particular, expresado en todas las métricas de evaluación. Esto se atribuye a que la forma matricial del cumulante y el biespectro contiene información posicional, que pueden ser interpretadas como imágenes. Las CNN están diseñadas para clasificar imágenes, a diferencia de las MLP que consideran a las matrices de entradas como un conjunto de datos descorrelacionados. Debido a los buenos resultados, se concluye que las CNN evitan de forma superadora el complejo y artesanal proceso de caracterizar las señales EMG.

Los resultados obtenidos son superiores comparativamente a los obtenidos previamente con el esquema clásico basado en características y con el modelo VGGNet. Esto último es de esperar debido a que la naturaleza de las imágenes de entrenamiento de VGGnet son distintas a las presentadas aquí. Se hace énfasis en que se usa un clasificador para todos los sujetos (versus el esquema usuario dependiente de los trabajos previos). Al analizar las mediciones de evaluación que se presenta cuando se usa un canal de EMG, en vez de los cuatro planteados originalmente, se obtienen resultados aceptables.

REFERENCIAS

- [1] L. H. Ting and J. L. McKay, "Neuromechanics of muscle synergies for posture and movement," *Current Opinion in Neurobiology*, vol. 17, no. 6, pp. 622–628, 2007.
- [2] C. J. De Luca, "Physiology and mathematics of myoelectric signals," *IEEE Transactions on Biomedical Engineering*, no. 6, pp. 313–325, 1979.
- [3] C. De Luca, "Surface electromyography: Detection and recording," *DelSys Incorporated*, vol. 10, no. 2, pp. 1–10, 2002.
- [4] M. Alom, T. Taha, C. Yakopcic, S. Westberg, P. Sidike, M. Nasrin, M. Hasan, B. Van Essen, A. Awwal, and V. Asari, "A State-of-the-Art Survey on Deep Learning Theory and Architectures," *Electronics*, vol. 8, no. 3, 2019.
- [5] T. Subba and T. S. Chingtham, "A Survey: EMG Signal-Based Controller for Human–Computer Interaction," in *Advances in Communication, Cloud, and Big Data*. Springer, 2019, pp. 117–125.
- [6] M. A. Oskoei and H. Hu, "Myoelectric control systems: A survey," *Biomedical Signal Processing and Control*, vol. 2, no. 4, pp. 275–294, 2007.
- [7] D. Roman-Liu and M. Konarska, "Characteristics of power spectrum density function of emg during muscle contraction below 30% mvc," *Journal of Electromyography and Kinesiology*, vol. 19, no. 5, pp. 864–874, 2009.
- [8] X. Zhang, Y. Wang, and R. P. Han, "Wavelet transform theory and its application in emg signal processing," in *2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery*, vol. 5. IEEE, 2010, pp. 2234–2238.
- [9] P. K. Artemiadis and K. J. Kyriakopoulos, "An emg-based robot control scheme robust to time-varying emg signal features," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 3, pp. 582–588, 2010.
- [10] X. Chen, X. Zhu, and D. Zhang, "A discriminant bispectrum feature for surface electromyogram signal classification," *Medical Engineering & Physics*, vol. 32, no. 2, pp. 126–135, 2010.
- [11] N. Sezgin, "Analysis of emg signals in aggressive and normal activities by using higher-order spectra," *The Scientific World Journal*, vol. 2012, 2012.
- [12] E. C. Orosco, N. M. Lopez, and F. di Sciascio, "Bispectrum-based features classification for myoelectric control," *Biomedical Signal Processing and Control*, vol. 8, no. 2, pp. 153–168, 2013.
- [13] F. Ayachi, S. Boudaoud, and C. Marque, "Evaluation of muscle force classification using shape analysis of the sEMG probability density function: a simulation study," *Medical & Biological Engineering & Computing*, vol. 52, no. 8, pp. 673–684, 2014.
- [14] E. Orosco, P. Diez, E. Laciár, V. Mut, C. Soria, and F. di Sciascio, "On the use of high-order cumulant and bispectrum for muscular-activity detection," *Biomedical Signal Processing and Control*, vol. 18, pp. 325–333, 2015.
- [15] E. C. Orosco and F. di Sciascio, "Muscular synergy classification and myoelectric control using high-order cross-cumulants," *Neural Computing and Applications*, vol. 28, no. 10, pp. 2979–2993, 2017.
- [16] C. L. Nikias and J. M. Mendel, "Signal processing with higher-order spectra," *IEEE Signal Processing Magazine*, vol. 10, no. 3, pp. 10–37, 1993.
- [17] A. Swami, C. Mendel, Jerry M, and C. L. Nikias, *Higher-Order Spectral Analysis Toolbox: For Use with MatLab*. MathWorks Incorporated, 1998.
- [18] D. R. Brillinger, *Time Series, Data Analysis and Theory*, siam ed. Holden Day Inc., 2001.
- [19] M. G. Kendall, *The Advanced Theory of Statistics I*, 2nd ed. Charles Griffin & Company Limited, 1945.
- [20] G. R. Naik, S. E. Selvan, M. Gobbo, A. Acharyya, and H. T. Nguyen, "Principal component analysis applied to surface electromyography: a comprehensive review," *IEEE Access*, vol. 4, pp. 4025–4037, 2016.
- [21] G. R. Naik, A. H. Al-Timemy, and H. T. Nguyen, "Transradial amputee gesture classification using an optimal number of sEMG sensors: an approach using ICA clustering," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 8, pp. 837–846, 2016.
- [22] M. C. Tresch, V. C. Cheung, and A. d'Avella, "Matrix factorization algorithms for the identification of muscle synergies: Evaluation on simulated and experimental data sets," *Journal of Neurophysiology*, vol. 95, no. 4, pp. 2199–2212, 2006.
- [23] S. A. Chvatal and L. H. Ting, "Voluntary and reactive recruitment of locomotor muscle synergies during perturbed walking," *The Journal of Neuroscience*, vol. 32, no. 35, pp. 12237–12250, 2012.
- [24] J. Roh, W. Z. Rymer, and R. F. Beer, "Robustness of muscle synergies underlying three-dimensional force generation at the hand in healthy humans," *Journal of Neurophysiology*, vol. 107, no. 8, pp. 2123–2142, 2012.
- [25] G. R. Naik and H. T. Nguyen, "Nonnegative matrix factorization for the identification of emg finger movements: evaluation using matrix analysis," *IEEE journal of biomedical and health informatics*, vol. 19, no. 2, pp. 478–485, 2015.
- [26] E. Orosco, N. López, C. Soria, and F. di Sciascio, "Surface electromyogram signals classification based on bispectrum," in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, 2010, pp. 4610–4613.
- [27] D. Buongiorno, G. D. Cascarano, A. Brunetti, I. De Feudis, and V. Bevilacqua, "A survey on deep learning in electromyographic signal analysis," in *International Conference on Intelligent Computing*. Springer, 2019, pp. 751–761.
- [28] A. Ameri, M. A. Akhaee, E. Scheme, and K. Englehart, "Real-time, simultaneous myoelectric control using a convolutional neural network," *PLoS one*, vol. 13, no. 9, p. e0203835, 2018.
- [29] R. Laezza, "Deep Neural Networks for Myoelectric Pattern Recognition An Implementation for Multifunctional Control," Ph.D. dissertation, Master Thesis, Chalmers University of Technology, Gothenburg, Sweden, 2018.
- [30] M. Zia ur Rehman, A. Waris, S. Gilani, M. Jochumsen, I. Niazi, M. Jamil, D. Farina, and E. Kamavuako, "Multiday EMG-based classification of hand motions with deep learning techniques," *Sensors*, vol. 18, no. 8, p. 2497, 2018.
- [31] H. M. A. Alaskar, "Deep Learning of EMG Time–Frequency Representations for Identifying Normal and Aggressive Actions," 2018.
- [32] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*. John Wiley and Sons, 2012.
- [33] C. L. Nikias and M. R. Raghuvver, "Bispectrum estimation: A digital signal processing framework," *Proceedings of the IEEE*, vol. 75, no. 7, pp. 869–891, 1987.
- [34] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [35] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [36] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, pp. 1345–1359, 2009.