

# DETECCIÓN DE AGARRE DE OBJETOS DESCONOCIDOS CON SENSOR VISUAL-TÁCTIL

Julio Castaño-Amorós, Pablo Gil, Ines Fernández, Santiago Puente  
 Grupo de Automática, Robótica y Visión Artificial, Universidad de Alicante  
 julio.ca@ua.es, pablo.gil@ua.es, ifs23@alu.ua.es, santiago.puente@ua.es

## Resumen

*La manipulación robótica sigue siendo un problema no resuelto. Implica muchos aspectos complejos como la percepción táctil de una amplia variedad de objetos y materiales, control de agarre para planificar la postura de la mano robótica, etc. La mayoría de los trabajos anteriores sobre este tema han estado utilizando sensores caros. Este hecho dificulta la aplicación en la industria. En este trabajo, se propone un sistema de detección de agarre mediante un sensor táctil de tecnología de imagen y bajo coste, conocido como DIGIT. El método desarrollado basado en redes convolucionales profundas es capaz de detectar contacto o no contacto, con precisiones superiores al 95%. El sistema ha sido entrenado y testado con una base de datos propia de más de 16000 imágenes procedentes de agarres de diferentes objetos, empleando distintas unidades de DIGIT. El método de detección forma parte de un controlador de agarre para una pinza ROBOTIQ 2F-140.*

**Palabras clave:** Percepción táctil, agarre robótico, sensor DIGIT, Redes Neuronales Convolucionales

## 1. INTRODUCCIÓN

La percepción táctil se está convirtiendo en esencial en cualquier tarea de manipulación robótica [8]. Los datos procedentes de sensores táctiles se han usado con éxito para obtener información del objeto que se está manipulando, como forma, rigidez o textura del material entre otras [9] con el objetivo de aplicarlo a tareas de reconocimiento de objetos [14]. También, es frecuente usar las señales táctiles para adaptar la fuerza con la que se agarran los objetos [4] cuando se manipulan o para planificar la pose del robot para asegurar un mejor agarre [2], todo ello con el objetivo de evitar que el objeto se escurra y se caiga.

Recientemente, han proliferado las aproximaciones basadas en técnicas de aprendizaje automático y aprendizaje profundo que tratan de resolver parcial o totalmente el problema de conseguir agarres robóticos estables de objetos [3], [1] y [17].

A pesar de la versatilidad que proporcionan las técnicas de aprendizaje, la gran mayoría de aproximaciones conocidas son muy dependientes de la tecnología táctil empleada y del tipo de datos o señal que proporcionan este tipo de sensores. Los sensores táctiles pueden ser piezo-resistivos, capacitivos, ópticos, magnéticos, con transductores barométricos, etc. [9]. Esto hace que, hasta la fecha, se hayan propuesto arquitecturas neuronales muy diversas en función de la naturaleza de la señal, del rango de ésta, del tipo de sensor y, por lo tanto, de la tarea táctil que se pretende llevar a cabo [18].

En este trabajo, se presenta una primera aproximación para detectar agarre de objetos desconocidos, es decir, de los que no se dispone de información previa, y además se hace uso de un novedoso sensor visual-táctil de bajo coste, conocido como DIGIT [7].

El resto del artículo se organiza de la siguiente manera: primero se describe el sistema de agarre robótico utilizado, formado por sensores DIGIT montados en una pinza ROBOTIQ 2F-140. En segundo lugar, se presenta la metodología empleada para la detección de contacto utilizando aprendizaje profundo. Finalmente, se muestran los resultados obtenidos junto con los trabajos futuros.

## 2. SISTEMA DE AGARRE

### 2.1. SENSOR VISUAL-TÁCTIL

Si se desea que un sistema robótico autónomo adquiera el nivel de manipulación humana, se requiere obligatoriamente dotarlo de sensores táctiles y, además, que éstos ofrezcan capacidades inteligentes para trabajar en entornos con cierto desconocimiento. En este trabajo se ha optado por construir varias unidades de un sensor de bajo coste del tipo visual-táctil, que sea fácilmente replicable para que pueda ser instalado en una gran variedad de aplicaciones de agarre. Dentro de la tecnología visual-táctil destacan los sensores Gelsight [16], Gelslim [5] y de nueva creación DIGIT [7].

DIGIT es un sensor que está compuesto de un elastómero, una lámina acrílica, una cámara de

color PCB y LEDs. El elastómero puede ser transparente o reflectivo y si se desea puede incorporar marcadores. Estos componentes se ensamblan en una carcasa formada por varias piezas de plástico o resina que pueden fabricarse con impresoras 3D (Fig. 1 y Fig. 2). La cámara captura hasta 60 imágenes/s y a una resolución de 320x240 píxeles.

En este trabajo, se han construido tres unidades distintas de sensores DIGIT, que denotamos en la experimentación como A, B y C. Todas las unidades que han sido construidas se realizaron con elastómeros reflectivos y sin usar marcadores. También, se ha incorporado y probado con diferentes tipos de material para la fabricación de la carcasa (PLA y PTEG) que envuelve todas las partes. Por lo que las tres unidades no son exactamente iguales, ya que el sensor C fue impreso en PLA y PTEG, mientras que A y B solo en PETG. Por otro lado, la dureza del gel influye en la deformación del elastómero. Esto es algo a tener en consideración para dotar al sensor de más o menos sensibilidad a la medida de contacto.

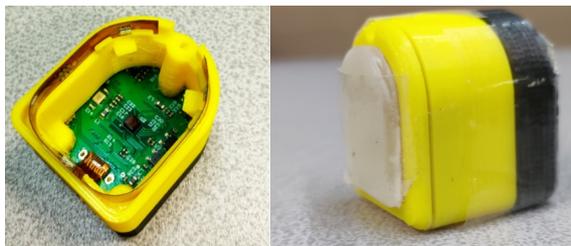


Figura 1: Sensor DIGIT.

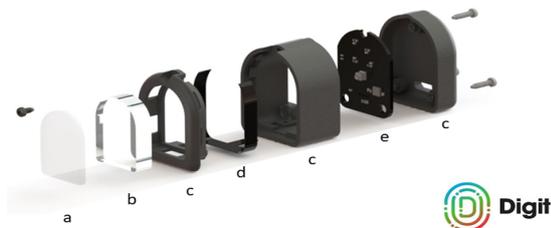


Figura 2: Partes del sensor DIGIT: a) Elastómero, b) Lámina acrílica, c) Carcasa 3D, d) LEDS, e) Cámara. Imagen extraída de [7].

## 2.2. PINZA ROBÓTICA

Para llevar a cabo los agarres, unidades de DIGIT se han montado en cada uno de los dos dedos de una pinza ROBOTIQ 2F-140 (Fig. 3). Este tipo de pinza tiene dos dedos articulados con dos articulaciones cada uno. Los dedos forman una única cadena cinemática por lo que se adapta a la forma del objeto durante el agarre y no pueden moverse de modo independiente. Este tipo de pinza puede ser empleada tanto para agarres internos como

externos. Los agarres internos consisten en aplicar fuerza desde el interior del objeto con la cara externa de los dedos, mientras que el agarre externo es el tradicional de aplicar fuerza con la cara interior de los dedos sobre la superficie externa del objeto. En este trabajo, los sensores se han montado de modo que su orientación permite obtener lecturas en configuración de agarre externo. Para montar estos sensores sobre los dedos de la pinza, se ha tenido que diseñar y construir una carcasa que permite integrar los sensores sin prácticamente restar espacio de trabajo de la pinza.

Una vez montados en la pinza, cada uno de los sensores por separado son conectados mediante USB, protocolo de comunicación serie, al mismo PC con el que se controlan los movimientos de la pinza. En este PC se emplea el software Robot Operative System (ROS) versión Kinetic Kame. En él se lleva a cabo tanto el control de la pinza como el procesamiento de las imágenes táctiles. Aunque el firmware de la pinza permite trabajar en un modo de operación en el que es posible fijar una fuerza final deseada para detectar cuando se produce contacto entre objeto y dedo y cuando no, este modo de operación no se ha empleado. El motivo de no emplearlo, es que este modo de funcionamiento no siempre permite una detección de contacto adecuada. Esto es porque es muy dependiente de la habilidad del sistema para medir el amperaje y compararlo con un amperaje de referencia definido a partir de la fuerza deseada, que a su vez es dependiente del tipo de objeto que se quiere agarrar. Si los materiales de los objetos que se pretende agarrar son distintos, estos valores de amperaje de referencia no tienen porque ser los mismos para todos, y habría que calibrar el sistema para ello. Razón por la que desde un punto de vista de incertidumbre, se ha preferido que sea un sensor táctil como DIGIT el que se encargue de esta tarea. DIGIT nos ha permitido trabajar en agarres de objetos distintos y desconocidos, sin tener un modelo matemático y sin necesidad de calibrado.

## 3. MÉTODO DE AGARRE

A continuación, se presenta el método empleado para agarre basado en la detección de contacto táctil. Como ya se ha mencionado, el método permite agarrar objetos distintos y desconocidos sin necesidad de disponer de complejos modelos matemáticos y sin requerir calibrado de la pinza/sensores. El método propuesto hace uso de técnicas de aprendizaje profundo. En concreto, una arquitectura neuronal con capas convolucionales entrenadas con una metodología de aprendizaje supervisado, que se comenta a continuación.

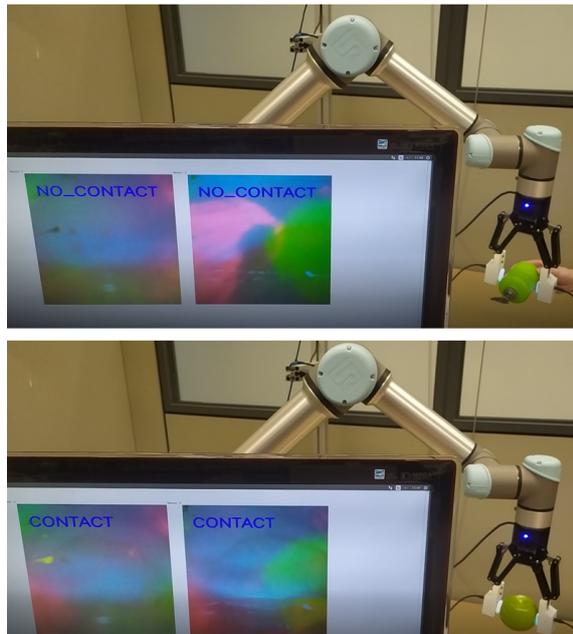


Figura 3: Sistema físico de agarre.

**3.1. ARQUITECTURA NEURONAL PROPUESTA**

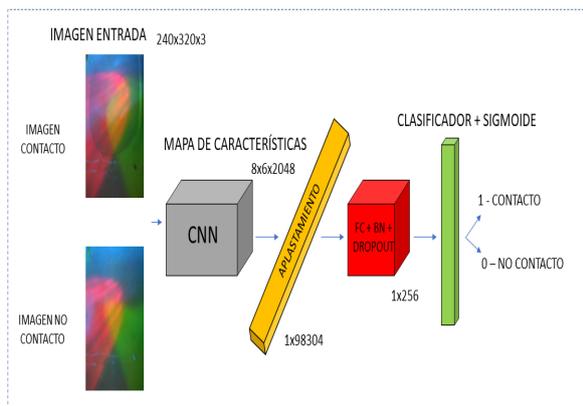


Figura 4: Arquitectura neuronal propuesta.

La arquitectura usada está compuesta de una estructura neuronal base compuesta por múltiples capas convolucionales (CL) y de agrupación (PL) que se emplea para la extracción de características de las imágenes táctiles que se toman como entrada. A ésta se le ha añadido una capa del tipo completamente conexas (FC) y un clasificador de una única neurona y una función de activación de tipo sigmoide, que nos permite una clasificación binaria de las imágenes táctiles a la entrada, para así detectar si hay o no hay contacto. Además, se ha aplicado un proceso de normalización por lotes (BN), así como se ha analizado la aplicación de un proceso de regularización empleando técnicas de "dropout" para evitar y/o reducir el sobreajuste (Fig. 4).

Esta arquitectura neuronal ha sido diseñada para su aplicación en el controlador propuesto en la Fig. 5. Con este controlador se pretende realizar una realimentación táctil basada en imagen, para garantizar un agarre estable.

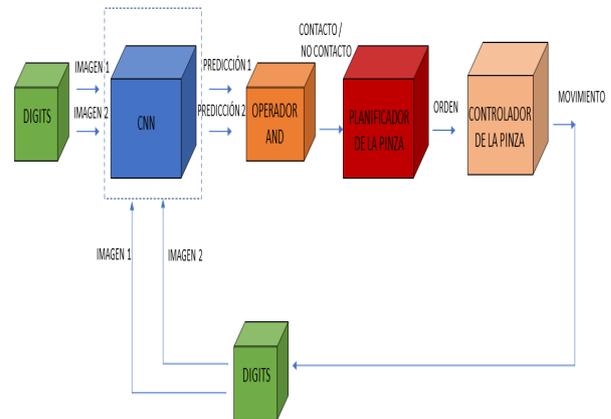


Figura 5: Controlador propuesto.

En trabajos previos como [17], en los que se empleaban imágenes táctiles sintéticas obtenidas a partir de los valores de los electrodos de sensores BIOTAC SP, era suficiente con emplear una red convolucional sencilla, la cual se limitaba a 3 o 4 capas convolucionales. Sin embargo, la complejidad de las imágenes táctiles obtenidas con DIGIT ha obligado a estudiar y analizar numerosos modelos neuronales profundos de clasificación, con un número mucho mayor de capas, para emplearlas como estructura neuronal base. Los modelos analizados pueden alcanzar varios cientos de neuronas por capa, y más de cincuenta capas en algunos casos. En concreto, en este trabajo se han plasmado algunos de los resultados obtenidos empleando como estructura convolucional base, modelos como VGG-19 [12], InceptionV3 [13] o MobileNetV2 [11]. Estos modelos, solo pretenden ser una muestra de los muchos que hay en el estado del arte, y que nos permite analizar la influencia de la profundidad, número de neuronas, el tiempo de entrenamiento o la precisión en la inferencia cuando se trabaja con imágenes táctiles de compleja interpretación como son las de DIGIT. A continuación, se describen brevemente las características de estos modelos base:

*VGG19* está compuesta por 19 capas convolucionales con filtros de 3x3, capas de agrupamiento de tipo máximo y 2 capas completamente conectadas al final. La capa de clasificación que por defecto emplea VGG ha sido sustituida por una capa adicional completamente conectada y un clasificador binario como ya se ha mencionado anteriormente.

*InceptionV3* es más compleja que VGG19 y está

compuesta por 48 capas, aunque algunas de estas capas en realidad son bloques de capas, formadas por capas más sencillas. Así, la principal diferencia radica en que incorpora capas adicionales conocidos como bloques Inception. Estos bloques son capas convolucionales cuyo tamaño de filtro varía dinámicamente, para de este modo extraer características con diferentes niveles de detalle en cada uno de los bloques de filtrado.

*MobileNetV2*, a diferencia de las otras, emplea muchos menos parámetros y operaciones matemáticas, reduciendo los tiempo de entrenamiento e inferencia. Además, de las conocidas capas convolucionales en la primera y penúltima capa de la red, así como una capa completamente conectada en la última, añade capas intermedias conocidos como bloques residuales de cuello de botella. Cada uno de estos bloques realizan una operación de suma que combina los resultados antes y después de aplicar tres pequeños filtros convolucionales.

### 3.2. DATOS DE ENTRENAMIENTO

Para llevar a cabo el entrenamiento y detección ha sido necesario crear una base de datos de imágenes táctiles procedentes de los tres sensores táctiles construidos (A, B y C). Las imágenes táctiles son obtenidas de la interacción de la pinza con diferentes objetos desconocidos. Es decir, el sistema no tiene conocimiento a priori de la forma, material o tamaño de éstos. La idea es conseguir una respuesta del sistema generalista, que permita detectar contacto con superficies de objetos cuyo comportamiento táctil por rigidez o textura, por ejemplo, difiera bastante de unos a otros. En total se han empleado 16444 imágenes táctiles procedentes de 9 objetos distintos: 4 variedades de suela, pelota, manzana, envase, aerosol y frasco (Fig. 6). De las que 8222 son de interacción con contacto con los objetos (40 % A, 30 % B y 30 % C), y otras 8222 registran no contacto (43 % A, 25 % B y 32 % C).



Figura 6: Representación de objetos empleados

## 4. EXPERIMENTACION

Para llevar a cabo la fase de entrenamiento y validación se ha usado una plataforma NVIDIA DGX A100. Y la fase de inferencia se realiza en un PC convencional i5-8400 @ 2.8Ghz, 16 GiB DDR4 RAM. Toda la experimentación se ha desarrollado en Python 3.7.3 con librerías Keras 2.4.0 y Tensorflow 2.4.1.

### 4.1. ENTRENAMIENTO

En la fase de entrenamiento, se ha llevado a cabo un ajuste hiperparamétrico para cada una de las estructuras base neuronales con las que se ha testado nuestra arquitectura. El objetivo es obtener un modelo subóptimo para cada modelo base, capaz de aprender contacto desde imágenes táctiles.

El conjunto de datos presentado en la sección 3.2, se ha dividido en tres subconjuntos: 70 % de las muestras se han empleado para entrenar, 20 % para validar y 10 % para testear. Las muestras para cada subconjunto se escogieron aleatoriamente, asegurando que no hay muestras idénticas en ninguno de los subconjuntos.

Notar que además, se han aplicado técnicas de aumento de datos para incrementar el número de muestras en la etapa de entrenamiento, no así en la de validación y test. De este modo, también se ha conseguido que todos los conjuntos estén bien balanceados, es decir, recojan el mismo número de muestras con contacto y sin él. Mientras que conseguir imágenes táctiles distintas de contacto es una labor costosa temporalmente pero sencilla, registrar imágenes distintas sin contacto no es fácil porque en ausencia de contacto el estado del sensor generalmente no cambia o varía muy poco. El proceso de aumento de datos ha consistido en transformaciones de ligeros zoom, rotaciones y operaciones de volteo.

En esta fase, se han probado distintos mecanismos de entrenamiento, desde entrenar la arquitectura desde cero, a emplear modelos preentrenados para resolver otros problemas de clasificación. Los mejores resultados se han obtenido a partir de modelos base preentrenados con un conjunto de datos completamente distinto (es decir no procedentes de sensores visual-táctiles), procedente de *ImageNet*, conjunto de imágenes RGB que se emplean para clasificar escenas de objetos [10] (ver Cuadro 1). En concreto, en el caso *VGG19*, el mejor resultado se obtuvo congelando los pesos de las 6 primeras capas y con reajuste del resto de capas. Se ha empleado tamaño de lote de 32, función de pérdida de entropía cruzada y optimizador de *Adam* con tasa de aprendizaje fijada en  $1e^{-6}$ , así como dropout de 0,5. Para el caso de *Incep-*

*tionV3*, se han congelado las 249 primeras capas y reajustado el resto. Todos los parámetros respecto a VGG se mantienen a excepción del optimizador empleado que ahora es *RMSprop* con una tasa de aprendizaje de  $1e^{-5}$ , este es más rápido que *Adam* y diseñado para redes más densas. El *dropout* se ha aumentado hasta 0,6. Finalmente, para el caso de redes más ligeras como es el caso de *MobileNetV2*, se han congelado solo las primeras 4 capas. Se ha aumentado el tamaño de lote hasta 64, y el resto de configuración es la misma que para *Inception*, aunque se ha prescindido de *dropout*. Mientras que *VGG19* ha requerido 100 épocas para aprender, el resto lo ha hecho en 10.

Cuadro 1: Resultados de test de nuestra arquitectura con los modelos base optimizados en función de las métricas Acc (exactitud), P (precisión), R (exhaustividad) y Time (tiempo de inferencia)

CNN	Acc	P	R	Time
<i>VGG</i>	99,9 %	99,8 %	100 %	140ms
<i>Inception</i>	99,7 %	99,5 %	100 %	90ms
<i>MobileNet</i>	98,1 %	97,3 %	98,9 %	70ms

El Cuadro 1 muestra los mejores resultados de la arquitectura presentada en la Fig. 4, empleando como CNN extractora de características, distintos modelos base. La arquitectura en todos los casos se testeó con el 10 % de las muestras reservadas para los tres sensores (A,B,C), siendo el promedio de éxito el que se muestra en la tabla.

#### 4.2. DETECCIÓN Y GENERALIZACIÓN

Finalmente, la última etapa de la experimentación ha consistido en determinar cómo de bien, el método propuesto, detecta contacto de agarre (Fig. 7) y además, si es capaz de detectar aún cambiando la unidad del sensor. En primer lugar, se pretende ver qué éxito de detección se obtiene cuando una misma unidad de DIGIT es empleada tanto para optimizar la arquitectura neuronal como para testear el agarre de distintos objetos. En segundo lugar, hasta que punto el éxito de detección se ve comprometido cuando se sustituyen las unidades DIGIT de la pinza por otras, por ejemplo como consecuencia de una rotura del elastómero. En este caso, la idea es seguir trabajando con el mismo modelo neuronal optimizado para la unidad estropeada, y sin necesidad de reajuste hiperparamétrico, conseguir que el sistema de detección se pueda seguir empleando para una nueva unidad DIGIT de sustitución. Este tipo de experimentación también es de mucho interés, aunque no se produzca rotura, cuando se quiere transferir resultados

a aplicaciones reales, por ejemplo reutilizando el código sin etapas de reajuste, si se quiere replicar el sistema en distintos robots.

En primera instancia, se ha optimizado la arquitectura propuesta con datos del sensor A (6186 muestras) y se ha testeado primero con muestras de agarre distintas de A (687), y después con muestras de un sensor B (436). Lo mismo se ha repetido para el sensor B. Esto es, se ha optimizado el sistema para datos de B (4367) y se ha testeado con nuevas muestras de B (436), y por separado con únicamente muestras de una unidad distinta, es decir, A (687).

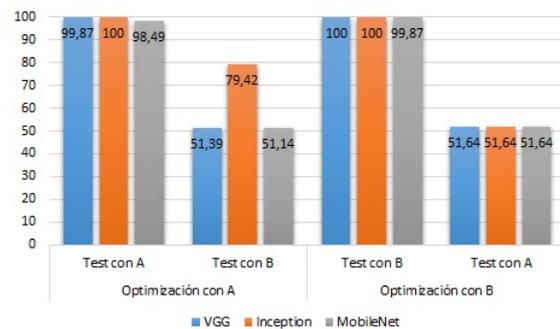


Figura 7: Resultados de detección I

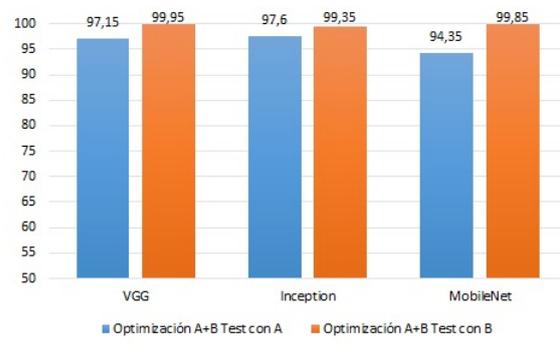


Figura 8: Resultados de detección II

Como se observa en la Fig. 7 el sistema alcanza tasas muy altas de detección  $\geq 98,5\%$  cuando se entrena con imágenes de la misma unidad del sensor que luego se instala en la pinza. Sin embargo, a la arquitectura propuesta le cuesta mucho generalizar (mejor de los casos 79 %). Es decir, funcionar correctamente cuando el sistema entrenado con una unidad es probado con una unidad de sensor distinta. Y esto se produce incluso aunque el modelo base sea distinto.

De aquí, se deduce que se requiere optimizar la arquitectura con variabilidad de muestras procedentes de distintas unidades de fabricación de DIGIT. Por lo tanto, en segunda instancia, se ha aumentado la variabilidad de muestras mezclando

en la optimización, muestras de distintas unidades, como son A y B (6186, 4367), y se han hecho tests con muestras nuevas de ambas unidades indistintamente, tal y como se muestra en la Fig. 8, probándose una importante mejoría en la detección  $\geq 94,3\%$ . Esto nos sugiere que es posible aumentar bastante la tasa de éxito de generalización si aumentamos la variabilidad de imágenes de entrenamiento, incluyendo otras procedentes de distintas unidades del sensor.

Además como ya se observó en el Cuadro 1, optimizando el modelo con los tres sensores construidos (A, B y C), y detectando agarre con cada uno de ellos indistintamente, todavía se obtienen mejoras a las presentadas en la Fig. 8. De ahí, es posible deducir que añadir muestras de diferentes unidades DIGIT mejora los resultados y puede ayudar a generalizar mejor.

## 5. CONCLUSIONES

En este trabajo, se ha implementado un sistema para detección de agarre basado en imagen táctil, la cual es generada por un novedoso sensor táctil-visual, conocido como DIGIT, de construcción no estandarizada. El método de detección, que está basado en redes neuronales convolucionales, muestra que es posible alcanzar una correcta detección con altas tasas de éxito, por encima del 95 % en la mayoría de los casos. No obstante, se hace difícil lograr una buena generalización entre diferentes unidades de sensor, principalmente como consecuencia de la fabricación artesanal que tienen. En cualquier caso, y a la vista de la experimentación, se cree que sería posible incrementando el número de unidades de fabricación optimizar el software para evitar los reentrenamientos y la dependencia de la unidad de sensor usada.

En la actualidad, se está trabajando en la mejora de la generalización, así como en la implementación de un controlador de agarre que haciendo uso del sistema de detección de agarre presentado, sea capaz de adaptar en tiempo real, la orientación de la pinza y la apertura de ésta durante la manipulación de objetos desconocidos, de los que no se tiene información a priori de éstos. Los resultados de este trabajo se pueden observar en el siguiente vídeo: <https://www.youtube.com/watch?v=TxoR9Xm1pcl>.

### Agradecimientos

Este trabajo ha sido financiado con FEDER a través del proyecto europeo COMMANDIA (SOE2/P1/F0638) de la convocatoria Interreg-V Sudoe. Además, se ha hecho uso de instalaciones de computación DGX-A100 adquiridas con una ayuda IDIFEDER/2020/003 del gobierno regional de

la Generalitat Valenciana.

## English summary

### GRASPING DETECTION OF UNKNOWN OBJECTS WITH VISUAL-TACTILE SENSOR)

#### Abstract

*Robotic manipulation is still a challenge. It involves many complex aspects such as tactile perception of a wide variety of objects and materials, grip control to plan robotic hand posture, etc. Most of the previous work used expensive sensors for tactile perception tasks. This fact implies difficulty in transferring application results to industry. In this work, a grip detection system is proposed. It uses DIGIT sensors based on low-cost image technology. The method developed, which is based on deep Convolutional Neural Networks (CNN), is capable of detecting contact or non-contact, with success rates greater than 95 %. The system has been trained and tested on our own dataset, composed of more than 16,000 images from different object grasping, also using several DIGIT units. The detection method is part of a grip controller used with a ROBOTIQ 2F-140 gripper.*

**Keywords:** Tactile perception, Robotic grasping, DIGIT sensor, Convolutional Neural Networks.

## Referencias

- [1] Bekiroglu, Y., Laaksonen, J., Jorgensen, J.A., Kyrki, V., Kragic, D., (2018) "Assessing grasp stability based on learning and haptic data", *IEEE Trans.on Robotics* 27(3), pp. 616-629. Doi: 10.1109/TRO.2011.2132870
- [2] Calandra, R., Owens, A., Jayaraman, D., Lin, J., Yuan, W., Malik, J., Adelson, E., Levine, S., (2018) "More than a feeling: Learning to grasp and regrasp using vision and touch", *IEEE Robotics and Automation Letters* 3(4), pp. 3300-3307. Doi: 10.1109/LRA.2018.285277
- [3] Cockburn, D., Roberge, J., Le, T., Masluczyk, A., Duchaine, V. (2017), "Grasp stability assessment through unsupervised feature learning of tactile images", *IEEE Int. Conf. on Robotics and Automation (ICRA)*, pp. 2238-2244. Doi: 10.1109/ICRA.2017.7989257
- [4] Delgado, A., Jara, C., Torres, F. (2017), "In-hand recognition and manipulation of

- elastic objects using servo-tactile control strategy”, *Robotics and Computer-Integrated Manufacturing* 48, pp. 102-112. Doi: 10.1016/j.rcim.2017.03.002
- [5] Donlon, E., Dong, S., Liu, M., Li, J., Adelson, E., Rodriguez, A. (2018), ”Gelslim: A high-resolution, compact, robust, and calibrated tactile-sensing finger”, *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, pp. 1927-1934. Doi: 10.1109/IROS.2018.8593661.
- [6] Kwiatkowski, J., Cockburn, D., Duchaine, V. (2017), ”Grasp stability assessment through the fusion of proprioception and tactile signals using convolutional neural networks”, *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, pp. 286-292. Doi: 10.1109/IROS.2017.8202170
- [7] Lambeta, M., Chou, P-W., Tian, S., Yang, B., Mallon, B., Most V.R., Stroud, D., Santos, R., Byagowi, A., Kammerer, G., Jayaraman, D., Calandra, R. (2020), ”DIGIT: A Novel Design for a Low-Cost Compact High-Resolution Tactile Sensor With Application to In-Hand Manipulation”, *IEEE Robotics and Automation Letters* 5(3), pp. 3838-3845. Doi: 10.1109/LRA.2020.2977257
- [8] Li, Q., Kroemer, O., Su, Z., Veiga, F.F., Kaboli, M., Ritter, H.J. (2020), ”A Review of Tactile Information: Perception and Action Through Touch”, *IEEE Trans. on Robotics* 36(6), pp. 1619-1634. Doi: 10.1109/TRO.2020.3003230
- [9] Luo, S., Bimbo, J., Dahiya, R., Liu, H. (2017), ”Robotic tactile perception of object properties: A review”, *Mechatronics* 48, pp. 54-67. Doi: 10.1016/j.mechatronics.2017.11.002
- [10] Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, C., Fei-Fei, L. (2015), ”ImageNet Large Scale Visual Recognition Challenge”, *Int. J. of Computer Vision (IJCV)*, pp. 211-252. Doi: 10.1007/s11263-015-0816-y
- [11] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.C. (2018), ”MobileNetV2: Inverted Residuals and Linear Bottlenecks”, *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 4510-4520. Doi: 10.1109/CVPR.2018.00474
- [12] Simonyan, K., Zisserman, A. (2015), ”Very Deep Convolutional Networks for Large-Scale Image Recognition”, *Int. Conf. on Learning Representations (ICLR)*. <https://arxiv.org/abs/1409.1556>
- [13] Szegedy, C., Vanhoucke, V., Ioffe S., Shlens J., Wojna, Z. (2015), ”Rethinking the Inception Architecture for Computer Vision”, *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)* pp. 2818-2826. Doi: 10.1109/CVPR.2016.308
- [14] Velasco, E., Zapata-Impata, B.S., Gil, P., Torres, F. (2020), ”Clasificación de objetos usando percepción bimodal de palpación única en acciones de agarre robótico”, *Revista Iberoamericana de Automática e Informática industrial* 17(1), pp. 44-55. Doi: 10.4995/riai.2019.1092310.3390/robotics8040085
- [15] Yi, Z., Zhang, Y., Peters, J. (2018), ”Biomimetic tactile sensors and signal processing with spike trains”, *A review. Sensors and Actuators A: Physical* 269, pp. 41-52. Doi: 10.1016/j.sna.2017.09.035
- [16] Yuan, W., Dong, S., Adelson, E. (2017), ”Gel-sight: High-resolution robot tactile sensors for estimating geometry and force”, *Sensors* 17(12), pp.2762. Doi: 10.3390/s17122762
- [17] Zapata-Impata, B.S., Gil, P., Torres, F. (2018), ”Non-Matrix Tactile Sensors: How Can Be Exploited Their Local Connectivity For Predicting Grasp Stability?”, *IEEE/RSJ Int. Conf. on Intelligent Robots And Systems (IROS). Workshop on Robotac: New Progress in Tactile Perception And Learning in Robotics*. <http://arxiv.org/abs/1809.05551>
- [18] Zapata-Impata, B.S., Gil, P., Torres, F. (2019), ”Tactile-Driven Grasp Stability and Slip Prediction”, *Robotics* 8(4), pp. 85:1-20. Doi: 10.3390/robotics8040085



© 2021 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution CC BY-NC-SA 4.0 license (<https://creativecommons.org/licenses/by-nc-sa/4.0/deed.es>).