



**HAL**  
open science

# Hand Posture Recognition Using Convolutional Neural Networks

Dennis Núñez Fernández, Bogdan Kwolek

► **To cite this version:**

Dennis Núñez Fernández, Bogdan Kwolek. Hand Posture Recognition Using Convolutional Neural Networks. LatinX in AI Research at ICML 2019, Jun 2019, Long Beach, United States. hal-02263891v2

**HAL Id: hal-02263891**

**<https://hal.science/hal-02263891v2>**

Submitted on 11 Sep 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Hand Posture Recognition Using Convolutional Neural Networks

Dennis Núñez Fernández<sup>1</sup>, Bogdan Kwolek<sup>2</sup>

<sup>1</sup> National University of Engineering, Peru

<sup>2</sup> AGH University of Science and Technology, Poland



## Introduction

### Hand Posture Recognition Problem

#### Objective:

- Reliable and fast hand posture recognition (to control Nao robot)

#### Approaches to the problem:

- Depth cameras (ToF, structured light, stereo-pair)
- RGB cameras
- Hand-crafted features
- Learned-features
- Convolutional Neural Networks

#### Problems:

- Difficulties in hand detection (in real-time)
- High dimensionality of the hand's kinematic configuration space
- Self-occlusions, variations in hand appearances
- Ill-posedness issues (the problem of inferring 3D coordinates from a single 2D observation is ill-posed)
- Camera geometry, scale difficulties, view-point
- Lack of dataset suitable for deep learning, see also [1]

## Dataset

### Dataset [3]:

- 6000 images of size 38x38
- 10 static postures
- 10 performers (various nationalities)
- Train/test: 5273/727 images (Kinect sensor)
- Test: 400 images (laptop - Dif. cam.)



Figure 3: Hand postures

## Our Approach

### Real-time hand posture recognition

- Hand detection, wrist position estimation
- CNN-based hand pose recognition
- Biologically inspired computer vision - Gabor filter

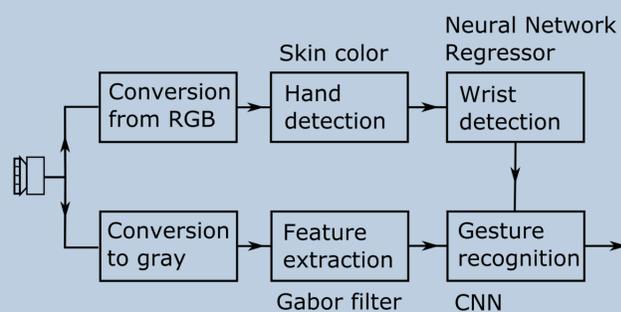


Figure 1: Block diagram of our method for hand posture recognition

## Overview of the Method

### Hand detection and wrist localization:

- Input: color images of size 640x480
- Statistical color model (RGB-H-CbCr) for skin detection
- Feature vector  $\mathbf{x}$  of size 38 on binary image
- Regression model:  $y = g(\mathbf{x}, \mathbf{w}) + \epsilon$ , where  $\epsilon$  is a random vector,  $\mathbf{w}$  are weights of a neural network with one hidden layer

### Convolutional neural network:

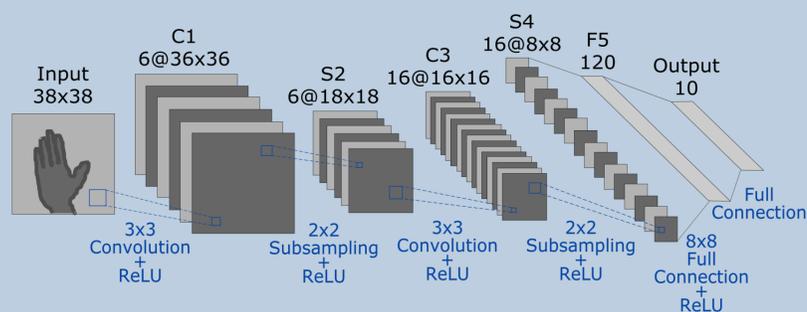


Figure 2: Architecture of convolutional neural network

### Gabor filter:

- Prefiltering images fed to CNN [2]
- Motivation: visual cortical simple cells were first shown by Hubel & Wiesel (Nobel Prize, 1981) to have response properties that resemble Gabor filters

## Experimental Results

	1	2	3	4	5	6	7	8	9	10	Avg.
Kinect	1.2	1.3	7.8	1.4	2.1	1.8	2.6	1.3	1.2	1.4	2.2
Dif. cam.	4.6	3.4	2.9	3.5	1.7	2.7	2.7	3.8	4.6	2.6	3.2

Table 1: Error of wrist location [pix] for classes 1-10

	Accuracy	Precision	Recall	F1 score
Kinect man.	0.950	0.945	0.960	0.949
Dif. cam. man.	0.928	0.931	0.937	0.934
Kinect aut.	0.905	0.906	0.932	0.906
Dif. cam. aut.	0.783	0.788	0.832	0.785

Table 2: Performance measures using CNN on raw gray images

	Accuracy	Precision	Recall	F1 score
Kinect man.	0.992	0.992	0.992	0.992
Dif. cam. man.	0.930	0.930	0.936	0.929
Kinect aut.	0.970	0.967	0.976	0.970
Dif. cam. aut.	0.868	0.868	0.882	0.866

Table 3: Performance measures using CNN with Gabor-based preprocessing

### Performance of the system:

	1	2	3	4	5	6	7	8	9	10	Avg.
Kinect	100	94	100	100	90	100	83	100	100	100	97
Dif. cam.	93	95	80	95	93	63	90	88	73	100	87

Table 4: Recognition accuracy [%] on the basis of wrist position estimated automatically

## Conclusions

- Promising results
- CNN operating on images prefiltered by Gabor achieves better classification performance
- Low computational costs

## References

1. P. Pisharady, M. Saerbeck: Recent methods and databases in vision-based hand gesture recognition. Comput. Vis. Image Underst., 2015.
2. B. Kwolek: Face detection using convolutional neural networks and Gabor filters, 2005.
3. <http://home.agh.edu.pl/~bkw/code/ciarp2017/>