ULPGC Universidad de Las Palmas de Gran Canaria | Escuela de Ingenierías Industriales y Civiles

Final Degree Project Curso 2022/2023
Grado en Ingeniería Electrónica Industrial y Automática GIEIA

Diseño de un Dispositivo Wearable para el Procesamiento de Datos Médicos Tabulares en Tiempo Real a través de una Red Neuronal Convolucional

Autor : Eduardo José Fernández Tovar Tutor: Eduardo Gregorio Quevedo Gutiérrez Co-Tutor: Antonio José Rodríguez Almeida

Introduction and objectives

Chronic Diseases (CD) account for 74% of worldwide. Designing efficient medical computer-aided system to improve early detection is key to diminish this. More specifically, wearable devices improve disease monitoring and detection. Convolutional Neural Networks (CNNs) are a widely used Deep Learning model especially useful with image input data, and research has been done to use it with tabular datasets. This Final Degree Project has three main objectives:

1. Develop a Python framework to study the use of Table to Image conversion (T2IC) algorithms on CD databases with CNNs.

2. Develop and profile a C based program to perform T2IC on data samples an infer using a CNN.

3. Study the feasibility of the hardware implementation for the developed C program.

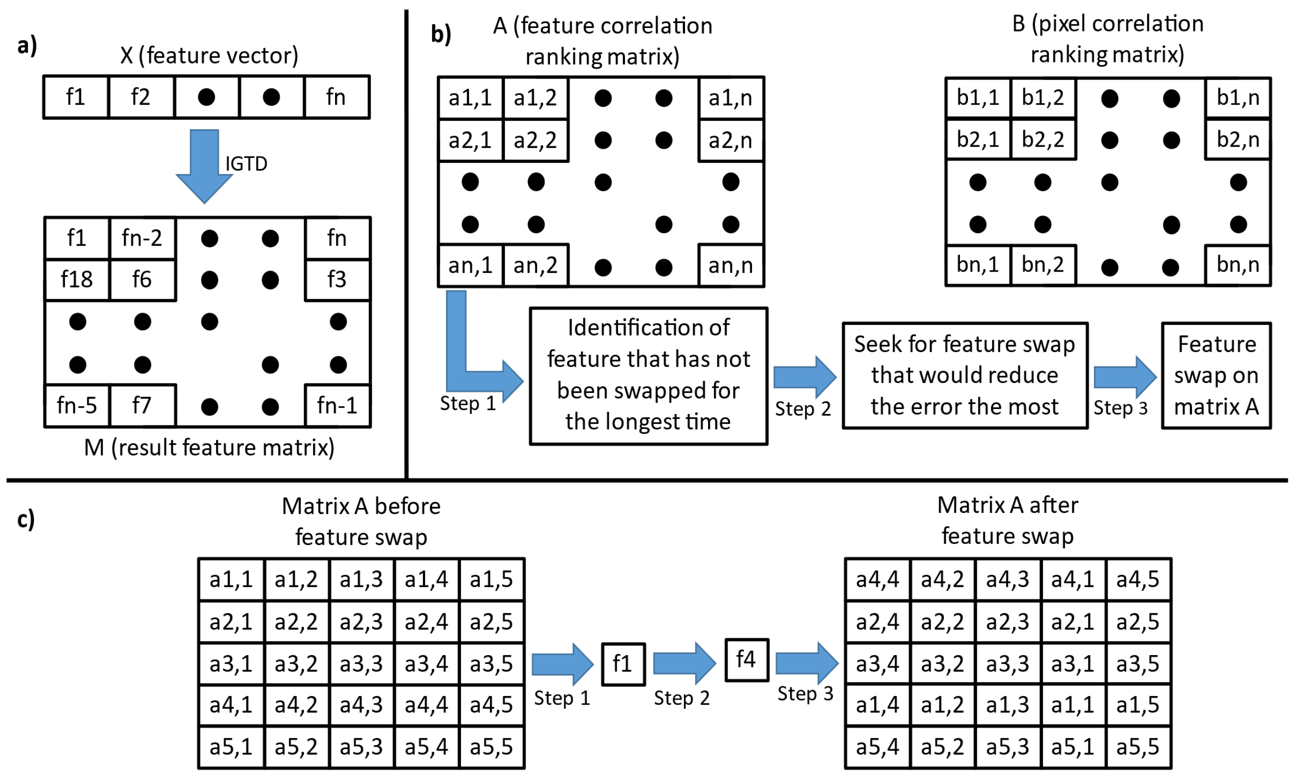

Figure 2: IGTD flow of conversion from tabular data to image

Results and conclusions

1. Databases with small number of features, like the Diabetes one, are unsuitable for this framework due to poor training results.

2. CNNs with DI images outperformed the CNNs models with IGTD images and the ML models used as reference.

3. Synthetic data generation techniques before applying T2IC algorithms improved CNNs performance.

4. We validated the program in C by comparing inference results in three steps of the process. The largest error observed was $1.6 \cdot 10^{-5}$ %, which was negligible.

5. Profiling revealed *conv* and *pad_image* were the most time consuming functions. We found out optimizing where *pad_image* is implemented reduced total time by 17.5 %

6. The program required 237 KB of storage and 15.20 KB of dynamic memory to be executed.

7. We conclude MIKROE-2502 is the embedded system suitable for most applications due to its low power consumption and memory specifications.

8. Successful implementation of the algorithm in an Arduino embedded system

9. Future work lines would include implementing the program in an embedded system of a wearable device and optimizing on-chip performance

Table 1: Best model performance

Metric	Accuracy	AUROC	Precision	Recall	F1
Value (%)	88.2	85.7	83.3	100.0	90.9

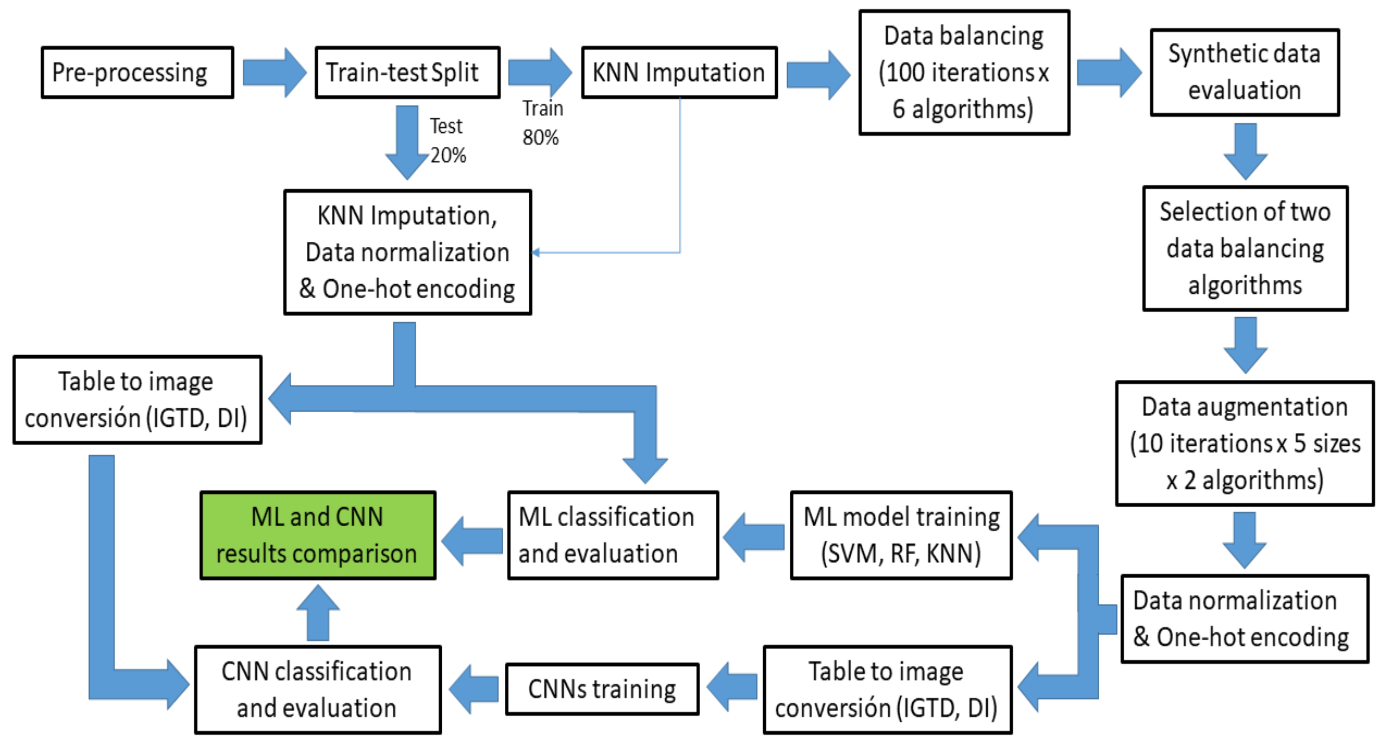

Figure 1: Python framework

Methodology

1. Application of synthetic data generation techniques to first balance the amount of control and cases of two databases, a Diabetes one and an Alzheimer one, and then augment the data, i.e. generate new synthetic samples. Two different T2IC were implemented in the framework, IGTD computes pixel correlation and feature correlation to place each features in the best pixel, while DI combines a similarity measuring technique (placing similar features together and dissimilar ones apart in a feature space) and a convex hull algorithm (to identify the smallest rectangle containing all the features) to generate images. ML models were trained on the tabular dataset and two CNNs were trained on the converted image dataset.

2. Conversion of python code of IGTD and DI to C and development of a code for inferring using a CNN. We validated and profiled the algorithm.

3. Study of several embedded systems and proposed three. Implementation of program in a hardware platform.

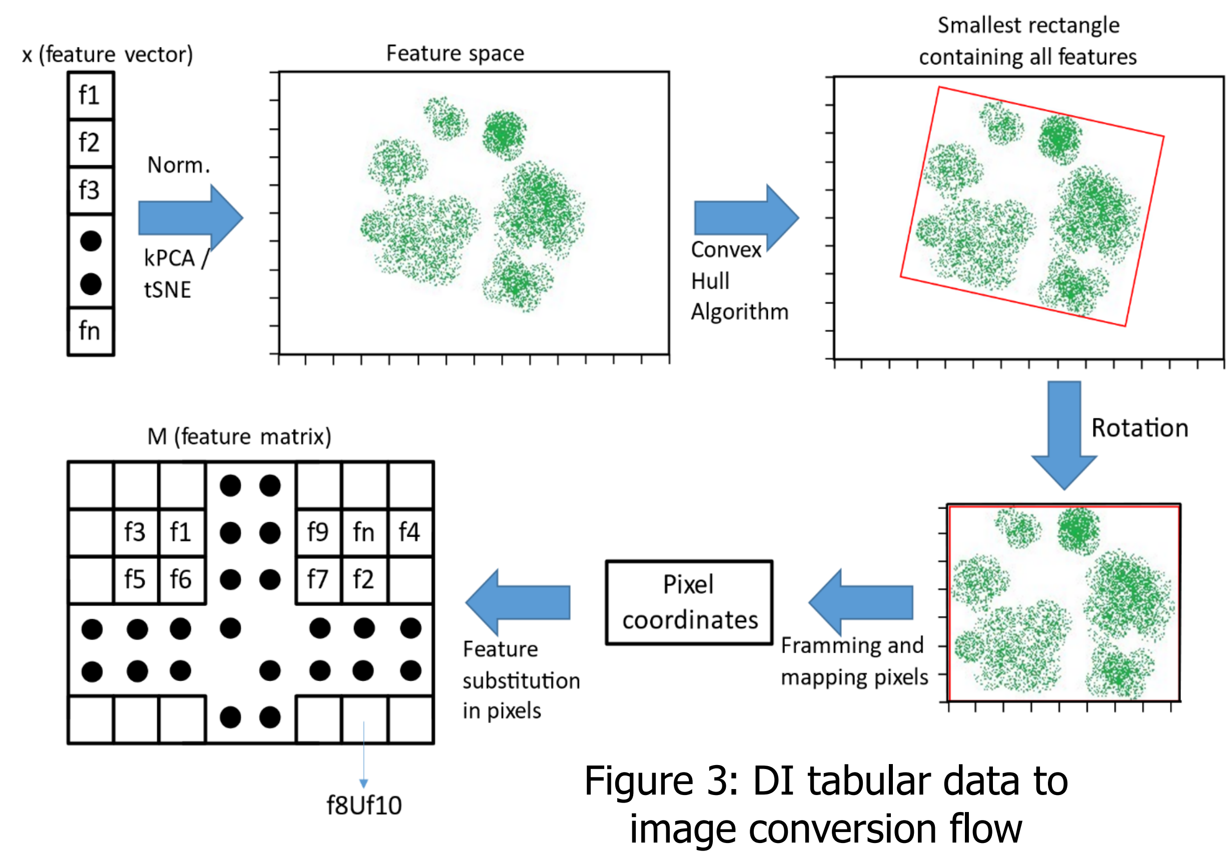

Figure 3: DI tabular data to image conversion flow

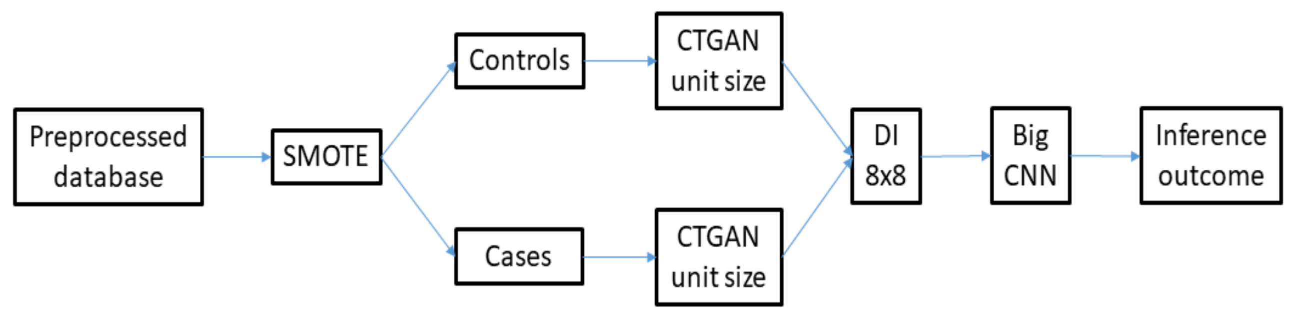

Figure 4: Best model workflow

Table 2: Arduino specifications

	Specifications
Processor	Cortex M0+
Max clock rate	133 MHz
Architecture	ARM (32 bits)
Flash main memory	16 MB
SRAM	264 KB
Communication	UART, SPI, I^2C, I^2S and PIO