

# MyDi application: towards automatic activity annotation of young patients with Type 1 diabetes

Lucia Migliorelli, Sara Moccia, Ismaela Avellino, Maria Chiara Fiorentino, Emanuele Frontoni

Department of Information Engineering, Università Politecnica delle Marche, 60131 Ancona, Italy

Email: {l.migliorelli@pm.univpm.it, s.moccia@univpm.it,  
ismaelaavellino@gmail.com, mariachiaraflorentino921@gmail.com,  
e.frontoni@univpm.it}

**Abstract**—Type I diabetes mellitus (T1DM) is a widespread metabolic disorder characterized by pancreatic insufficiency. People with T1DM require: a lifelong insulin injection, to constantly monitor glycemia and to take note of their activities. This continuous follow-up, especially at a very young age, may be challenging. Adolescents with T1DM may develop anxiety symptoms and depression which can lead to the loss of glycemic control. An assistive technology that automatizes the activity monitoring process could support these young patient in managing T1DM. The aim of this work is to present the MyDi framework which integrates a smart glycemic diary (for Android users), to automatically record and store patient’s activity via pictures and a deep-learning (DL)-based technology able to classify the activity performed by the patients (i.e., *meal* and *sport*) via picture analysis. The proposed approach was tested on two different datasets, the Insta-Dataset with 3498 pictures (also used for training and validating the DL model) and the MyDi-Dataset with 126 pictures, achieving very encouraging results in both cases ( $Prec_i = 1.0$ ,  $Rec_i = 1.0$ ,  $f1_i = 1.0$  with  $i \in C$ : [*meal*, *sport*]) prompting the possibility of translating this application in the T1DM monitoring process.

**Index Terms**—Type I Diabetes Mellitus, activity recognition, automatic activity annotation, picture analysis, deep learning

## I. INTRODUCTION

Type I diabetes mellitus (T1DM) is a widespread metabolic disorder: Cho et al. [1] identified a prevalence of T1DM in population aged 18-99 years of 8.4 % in 2017 and predicted to reach 9.9 % up to 2045. T1DM is characterized by an autoimmune deterioration of the pancreatic  $\beta$  cells, which culminates in a deficiency of insulin production with a consequent chronic elevation of blood glucose levels (hyperglycemia) [2]. Nowadays, to monitor the blood glucose levels, a popular solution is using subcutaneous continuous-monitoring systems, such as Dexcom<sup>®</sup>. Dexcom<sup>®</sup> measures glucose levels in the interstitial subcutaneous fluid and consists of: (i) a glucose sensor, usually implanted in the subcutaneous tissue of abdomen or upper glutes, (ii) a monitor to read and display the glucose level in the plasma (Fig. 1), (iii) a transmitter that allows the sensor and the monitor to communicate (via cable or wireless).

So far, no definitive cure for T1DM exists and the conventional treatment to control blood glucose level consists of insulin injection to simulate  $\beta$  cells action, but managing the right insulin injection, especially at a very young age, can be challenging [3]. To define a proper insulin dose to be injected,

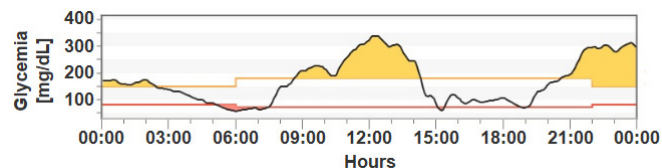


Fig. 1. An example of a daily glycemic curve from the Dexcom<sup>®</sup> sensor. Hyperglycemic (over the yellow line) and hypoglycemic (under the red lines) episodes are highlighted in yellow and red, respectively.

the patients are required to manually annotate the activity performed during the day (e.g., physical activity and meal ingestion) on a glycemic diary to possibly adapt the insulin therapy in case of hypo- or hyper-glycemic events (Fig. 1).

Keeping the glycemic diary updated is a tedious and time-consuming procedure, which often results in a poor or sporadic annotation [4]. Moreover, this everlasting follow-up strongly impacts patient’s quality of life and his/her psychological status. Adolescents with diabetes are at a material risk of developing psychiatric or eating disorders, as well as substance abuse; these issues can cause refusal of treatment and degradation of glycemic control [5].

An assistive technology that automatizes the monitoring process could support the young patient and his/her parents in controlling glucose metabolism and represent a powerful tool for telemedicine applications. In this scenario, some preliminary applications that link glucose levels to the activity performed by the patient have been proposed [6], but still require to update a glycemic diary [7].

To attenuate the tediousness of manual annotation, a possible solution could be making the activity annotation process more automatic via image-based diary records, which recently emerged as a novel method for activity assessment [8], [9]. Hence, in this work we propose the MyDi framework, which integrates: (i) the MyDi smartphone application, a smart glycemic diary to retrieve patients’ activity via pictures taken with their smartphone, and (ii) a deep-learning (DL)-based technology trained to automatically recognize the activity performed by the patients via images. In this paper, we focused on two macro activity classes, *meal* and *sport*, as these events are among the main causes of hyperglycemic and hypoglycemic episodes [2].

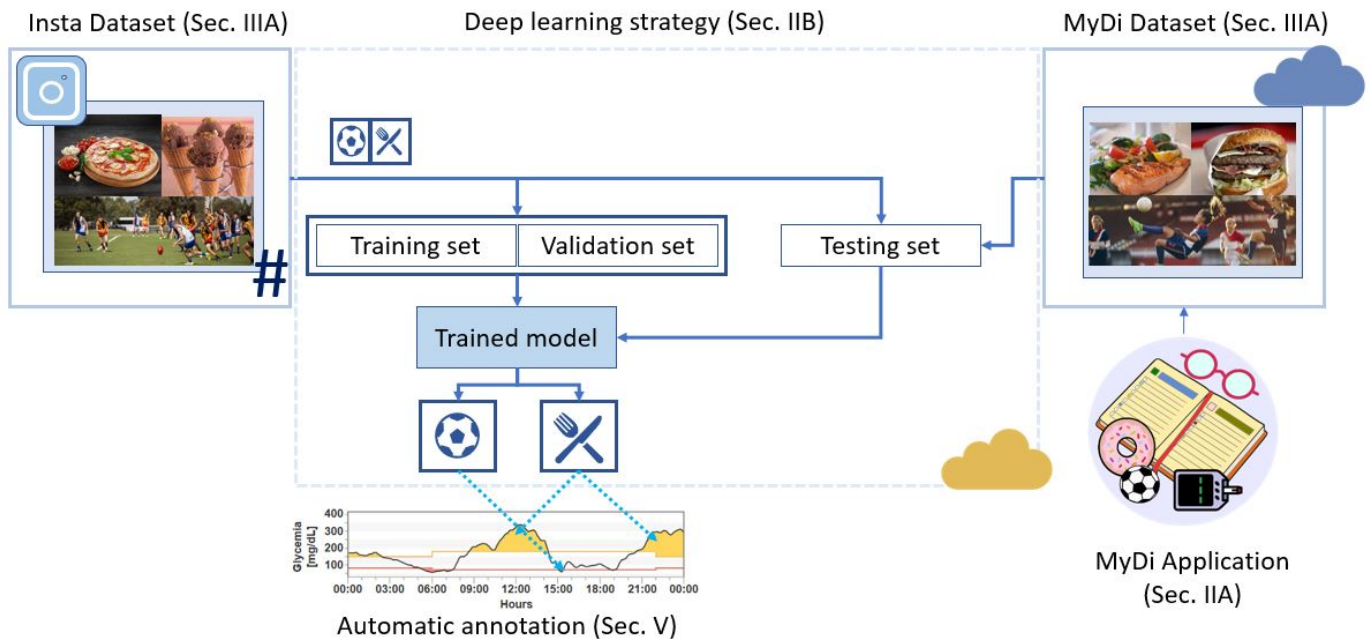


Fig. 2. Workflow of the proposed approach for activity recognition from pictures for diabetes monitoring in young patients.

This paper is organized as follows: Sec. II presents the MyDi app and the DL-based approach for classifying images. Sec. III, describes the experimental protocol for validation. The obtained results are presented in Sec. IV and discussed in Sec. V which also summarizes the main achievements and impact of the work.

## II. METHODS

This section describes the MyDi smartphone application (Sec. II-A) and the DL strategy to image-content retrieval (Sec. II-B). The workflow of the proposed approach is shown in Fig. 2.

### A. MyDi application

MyDi application, programmed in C Sharp and is available for Android users. It is designed as an alternative to the glycemic diary with manual annotation and allows the user to keep track of the activities performed by taking pictures and clicking icons in an interactive way. When the patient uses MyDi for the first time he/she completes a calendar with the planned activities on a weekly basis.

Based on pre-annotated activities, notifications appear to remind the patient to take pictures of his/her timetabled activity. The application will automatically keep track of the time at which the photo is taken and this time will then be traced back to the measurements of the Dexcom<sup>®</sup> sensor.

All the information is then stored on a cloud-based platform with a strong focus on data protection, security and anonymity.

### B. Deep learning strategy

In order to classify images from MyDi application, for understanding the activity performed by the patient (i.e., *meal*

or *sport*), a VGG16 convolutional neural network (CNN) was used (Fig. 3).

In the original VGG16 implementation, the input  $224 \times 224$  RGB image is processed through 13 convolutional (conv) layers for extracting image features. Each conv block has filters with a very small receptive field ( $3 \times 3$  pixels) and is followed by a rectified linear unit (ReLU) activation function.

Max pooling layers are inserted after 2 or 3 convolutional blocks. Pooling has the purpose of progressively reducing the spatial size of the feature map with the final goal of lowering the amount of parameters, reducing the computational complexity and consequently the risk of overfitting. Finally, 3 fully-connected layers - 4096 neurons in the two first layers and 1000 neurons in the last one - followed by a softmax layer are used to predict a probabilistic label map.

To fulfill our binary classification task, fine tuning was implemented. This procedure avoided training the network from scratch by uploading ImageNet<sup>1</sup> pre-trained weights and training a small mini network. To construct this mini network the first 4 conv blocks of VGG16 were frozen and the 3 fully connected layers were modified - 1024 neurons in the first layer, 512 neurons in the second and 2 neurons in the last one - and added to the remaining conv block (Fig. 3)

There are many different reasons to apply fine tuning : (i) the training time is drastically reduced, (ii) performance is improved as the model is trained on large scale dataset, (iii) lower and mid-level filters are not adapted specifically to a small dataset, not leading the model to overfit as the features extracted from ImageNet database are very generic [10].

<sup>1</sup><http://www.image-net.org/>

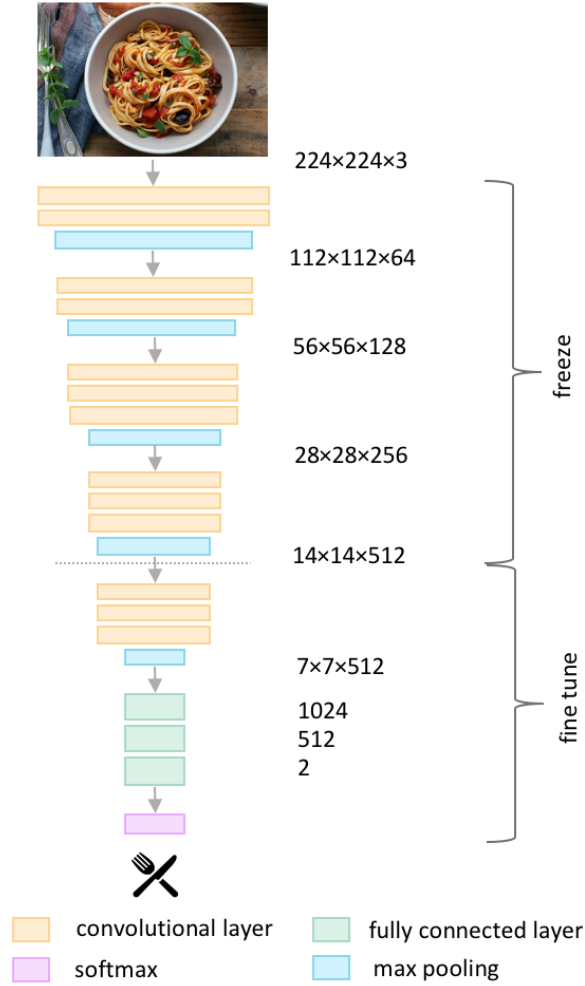


Fig. 3. VGG16 architecture with a focus on fine tuning. Fine tuning is implemented by freezing the first 4 convolutional (conv) blocks and constructing a mini network of 3 newly stated fully connected layers - 1024 neurons in the first layer, 512 neurons in the second and 2 neurons in the last one - added to the remaining conv block.

### III. EXPERIMENTAL PROTOCOL

#### A. Collected dataset

For fine-tuning VGG16, “digital footprints” from the popular social network *Instagram* were downloaded using the Python web crawler *Instaloader*, which was designed to retrieve images with the tag of interest [11]. A dataset (Insta-Dataset) of 3498 images was collected using varied hashtags, both on trend and specific ones, such as #breakfast, #eating, #yummy and #italianfood for *meal* and #volleyball, #footballpractice, #baseballmatch and #swimming for *sport*. The dataset was then divided into training (2000), validation (1000) and testing (498) images, which were all equally balanced between the two classes.

A further testing dataset (MyDi-Dataset) was collected exploiting the MyDi application. 10 volunteers used the app (after signing a written consent for privacy issues) throughout

TABLE I  
DATASET INFORMATION.

	Training	Validation	Testing
Insta-Dataset	2000	1000	498
MyDi-Dataset	-	-	126

a period of ten days simulating T1DM patients, uploading pictures of their meals and physical activities in order to assemble a collection of images as realistic as possible.

Datasets information is summarized in Table I.

#### B. Training settings

Before training the network, RGB images were resized to 224x224 pixels, which is the size of images in the Imagenet dataset.

Then VGG16 was fine-tuned employing stochastic gradient descend (SGD) as the optimizer, which helps the model avoid degenerate solutions [12], and using the binary cross-entropy as *Loss* function for 100 epochs. The batch size was 32 as default with a learning rate = 0.0001. The best model was then selected according to the highest accuracy (*Acc*) (Eq. 1) on the validation set.

$$Acc = \frac{\sum_k^i TP_k}{n}, i \in C : [meal, sport] \quad (1)$$

Where:  $C$  represents the classes set,  $TP_k$  represents the correctly classified sample and  $n$  is the total number of samples. To fine-tune VGG16, Keras library<sup>2</sup> was used.

All experiments were implemented on Google Colaboratory<sup>3</sup>: a free GPU cloud platform.

#### C. Performance metrics

To measure the performance of the classification CNN, we computed the classification Precision ( $Prec_i$ ) (Eq. 2), Recall ( $Rec_i$ ) (Eq. 3), and f1-score ( $f1_i$ ) (Eq. 4) for the  $i$ -th class, with  $i \in C : [meal, sport]$

$$Prec_i = \frac{TP_i}{TP_i + FP_i} \quad (2)$$

$$Rec_i = \frac{TP_i}{TP_i + FN_i} \quad (3)$$

$$f1_i = \frac{2 \times Prec_i \times Rec_i}{Prec_i + Rec_i} \quad (4)$$

where  $TP_i$ ,  $FP_i$ ,  $FN_i$  are the correctly classified samples, the false positives and the false negatives for the  $i$ -th class, respectively.

<sup>2</sup><https://keras.io/>

<sup>3</sup><https://colab.research.google.com/notebooks/welcome.ipynb#recent=true>

TABLE II

PERFORMANCE METRICS ACHIEVED ON THE *Instagram* TESTING DATASET. METRICS ARE REPORTED IN TERMS OF PRECISION ( $Prec_i$ ), RECALL ( $Rec_i$ ) AND F1-SCORE ( $f1_i$ ) WHERE  $i$  REPRESENTS THE  $i$ -TH CLASS.

	$Prec$	$Rec$	$f1$
<i>meal</i>	0.99	1.0	0.99
<i>sport</i>	1.0	0.99	0.99

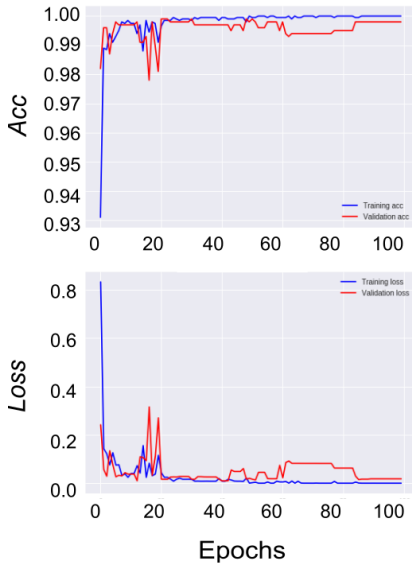


Fig. 4. Learning curves to evaluate the training (blue) and validation (red) Accuracy ( $Acc$ ) and Loss.

#### IV. RESULTS

Tab. II shows the results achieved by the network in terms of  $Prec_i$ ,  $Rec_i$  and  $f1_i$  (with  $i \in C: [meal, sport]$ ) for Insta-Dataset. MyDi-Dataset was used for testing too and the same metrics were calculated. In this latter case  $Prec_i$ ,  $Rec_i$  and  $f1_i$  were all equal to 1.

In Fig. 4 learning curves to evaluate the training and validation  $Acc$  and  $Loss$  are depicted. The best DL-model was the one with the highest  $Acc$  as reported in Fig. 4. Moreover, to prove that the model did not experience the overfitting phenomenon we also reported in the same Figure (Fig. 4) the  $Loss$  curve for the validation set with its decreasing trend.

#### V. DISCUSSION AND CONCLUSION

This work presented the MyDi framework which coupled the MyDi smartphone application with VGG16 CNN to support young adults with diabetes in automatizing the process of activities annotation. Hence, the system, did not force the patient in keeping a glycemic diary but directly classified his/her activity from pictures.

Two datasets were considered to test the network: the Insta-Dataset (which was also used for training and validating the network) and the MyDi-Dataset. As reported in Sec. IV, encouraging results were achieved for both datasets and the two classes considered (i.e., *meal* and *sport*) with peaks of 1.0 in all the 3 metrics ( $Prec_i$ ,  $Rec_i$ ,  $f1_i$ ). These results

suggested that MyDi could represent a valid solution for young patients with diabetes to monitor their pathology in an easy and motivating way, without having to manually record their activity on a diary.

However to give them the chance to improve the overall quality of their life it is acknowledged that further research is required. Among the possible future improvements the CNN could be trained to recognize various parameters and characteristics from the photo taken through MyDi and retrieve the amount of food eaten, of kilocalories consumed or of carbohydrates ingested or interpreting if the exercise is aerobic or anaerobic. Moreover the trained model could be uploaded on the new generation of smartphone with processors able to handle CNN to make online predictions. Other kinds of data could be integrated in the application (as in [13]): for example the ones coming from wearable sensors (e.g., Fitbit<sup>4</sup>) and other health applications (e.g., Android HiCare) that would add more precise information on the activities carried out by the patient and allow a more robust and continuous (data would be retrieved at different times during the entire day) estimation of the latter. Furthermore, directly correlating blood glucose variation detected with Dexcom<sup>®</sup> to the activity (*meal* and *sport*) performed by the patient, uploaded on MyDi, would be clinically relevant.

Organizing the entirety of these data could benefit clinic knowledge of T1DM, giving the chance to study young patients' behaviour and routines more in detail.

#### REFERENCES

- [1] N. Cho, J. Shaw, S. Karuranga, Y. Huang, J. da Rocha Fernandes, A. Ohlogge, and B. Malanda, "IDF diabetes atlas: global estimates of diabetes prevalence for 2017 and projections for 2045," *Diabetes Research and Clinical Practice*, vol. 138, pp. 271–281, 2018.
- [2] H. Baynes, "Classification, pathophysiology, diagnosis and management of diabetes mellitus," *Journal of Diabetes and Metabolism*, vol. 6, no. 5, pp. 1–9, 2015.
- [3] A. D. Association *et al.*, "12. children and adolescents: Standards of medical care in diabetes 2018," *Diabetes Care*, vol. 41, no. Supplement 1, pp. S126–S136, 2018.
- [4] D. S. Katz, B. A. Price, S. Holland, and N. S. Dalton, "Data, data everywhere, and still too hard to link: Insights from user interactions with diabetes apps," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 2018, p. 503.
- [5] M. Strand, A. Broström, and A. Haugstvedt, "Adolescents' perceptions of the transition process from parental management to self-management of type 1 diabetes," *Scandinavian Journal of Caring Sciences*, vol. 33, no. 1, pp. 128–135, 2019.
- [6] B. Cvetković, V. Janko, A. E. Romero, Ö. Kafalı, K. Stathis, and M. Luštrek, "Activity recognition for diabetic patients using a smartphone," *Journal of Medical Systems*, vol. 40, no. 12, p. 256, 2016.
- [7] J. Given, M. O'kane, B. Bunting, and V. Coates, "Comparing patient-generated blood glucose diary records with meter memory in diabetes: a systematic review," *Diabetic Medicine*, vol. 30, no. 8, pp. 901–913, 2013.
- [8] L. Ciabattini, F. Ferracuti, G. Lazzaro, L. Romeo, and F. Verdini, "Serious gaming approach for physical activity monitoring: A visual feedback based on quantitative evaluation," in *2016 IEEE 6th International Conference on Consumer Electronics-Berlin (ICCE-Berlin)*. IEEE, 2016, pp. 209–213.

<sup>4</sup><https://www.fitbit.com/it/home>

- [9] L. Ciabattoni, G. Foresi, A. Monteriù, D. P. Pagnotta, L. Romeo, L. Spalazzi, and A. De Cesare, "Complex activity recognition system based on cascade classifiers and wearable device data," in *2018 IEEE International Conference on Consumer Electronics (ICCE)*. IEEE, 2018, pp. 1–2.
- [10] Y.-X. Wang, D. Ramanan, and M. Hebert, "Growing a brain: Fine-tuning by increasing model capacity," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 2471–2480.
- [11] K. Han, Y. Jo, Y. Jeon, B. Kim, J. Song, and S.-W. Kim, "Photos don't have me, but how do you know me?: Analyzing and predicting users on instagram," in *Adjunct Publication of the 26th Conference on User Modeling, Adaptation and Personalization*. ACM, 2018, pp. 251–256.
- [12] S. Mandt, M. D. Hoffman, and D. M. Blei, "Stochastic gradient descent as approximate bayesian inference," *The Journal of Machine Learning Research*, vol. 18, no. 1, pp. 4873–4907, 2017.
- [13] L. Ciabattoni, F. Ferracuti, S. Longhi, L. Pepa, L. Romeo, and F. Verdini, "Real-time mental stress detection based on smartwatch," in *2017 IEEE International Conference on Consumer Electronics (ICCE)*. IEEE, 2017, pp. 110–111.