Cut Quality Estimation in Industrial Laser Cutting Machines: A Machine Learning Approach

Giorgio Santolini
University of Trento
giorgio.santolini94@gmail.com

Paolo Rota
University of Trento
paolo.rota@unitn.it

Davide Gandolfi
Adige S.p.A.
lasertech@adige.it

Paolo Bosetti
University of Trento
paolo.bosetti@unitn.it

Abstract

The use of machine learning models to improve industrial production quality is becoming more popular year after year. The main reason is the huge data availability and the impressive boost of performance of such methods achieved in the last decade. In this work we propose an adaptation of three well known machine learning algorithms to estimate the quality of cut in industrial laser cutting machines. The challenge here is to use a pool of multimodal parameters coming from different sensors and fuse them in order to detect the cutting status of the machine in a near-online modality. We analyze then generative and discriminative approaches based on Gaussian Mixture Models, Recurrent Neural Networks, and Convolutional Neural Networks in a supervised setting. Results are computed on a brand-new dataset that is freely available for reference.

1. Introduction

In recent years machine learning has attracted a lot of attention for many applications in industry spanning from predictive maintenance [15, 16], self-driving cars [23, 22] and healthcare [18, 21] to name a few. The main reasons for such popularity boost is due to the large and diverse amount of digital data that is now available and to the extraordinary development of the GPU computing that enables processing of data in parallel at high speed. In this regards, Neural Networks have played a big role, enhancing performance of machine vision algorithms to an impressive level [12]. Industries and research institutions are pushing on this theme, researching methods to efficiently fuse information coming from different modalities in order to produce a better inference [17, 10, 3, 5].

In this work we focus on a particular application of predictive process analysis, exploiting data generated by an advanced laser cutting machine (i.e. LT8.10 manufactured by Adige S.p.A. in Figure 1) aiming at detecting the cut interruption in this type of systems while cutting metal tubes or sheets. In this work we propose an extensive experimentation of different machine learning approaches to classify multiple process status indicators. Industrial laser cutting machines are built to perform fast, precise, and high quality cut of metal material. The possibility to cut different types of metal of multiple thicknesses using different cut technologies requires the machine manufacturer to produce a large amount of presets optimized for each combination of these properties. This procedure is named *characterization* and entails a highly time-consuming task. However, these properties are often not homogeneous in the material: for example, if we consider a metal tube obtained by folding a metal sheet and welding along its edges, it might show an uneven surface and thickness on the welding zone, often resulting in bad quality cuts or, in the worse case, a lack of continuity in the cut. Welding is not the only troublesome case; similar conditions might be produced by rust that locally changes the reflectance of the material, or by a wrong selection of cutting parameters by the operator.

Figure 1. Adige 3D laser cutting machine LT8.10. Courtesy of Adige S.P.A. - BLM Group

In order to close the loop and allow the machine to be able of automatically understanding when the cut has not been performed correctly is therefore of paramount importance to analyze the process ongoing in a near-real-time manner, i.e. with latency short enough to allow to take effective counter-measures.

Our goal is then to identify the loss of cut before the piece is welded and ruined in order to perform some corrective actions and avoid the stop of the production. In order to do so, we use process parameters that are known to, and commanded by, the machine such as laser power, gas pressure, cut speed, type of geometry being cut, type of material and type of gas. Together with these parameters we also have the opportunity to use the light collected through the optical fiber itself [19]. For the purpose of this work we have recorded the light intensity synchronously to the process parameters.

We have collected and labelled a large amount of data, cutting more than 20 pieces where each piece has about 20 geometries each for over 400 single example cuts. Each example is a time series of measures of the above mentioned parameters. Each geometry counts an average of about 90 thousands measures.

To tackle this problem we performed an evaluation of 3 different approaches: the first is a generative method where each class distribution has been modelled by a Mixture of Gaussians and the class decision is then sampled by the trained model. The other two methods are discriminative models based on deep neural networks. In particular, the first is based on a Long Short-Term Memory Network (LSTM) [8], and the second on a Fully Convolutional Neural Network [14] both properly adapted for the task in exam.

Our main contributions can be summarized in the following points:

- We have collected a dataset that to the best of our knowledge is the biggest dataset for automatic laser cutting finalized to the understanding of the quality of this process.
- We adapted and tested three well-known machine learning algorithms to the task, producing a thorough set of results.
- We proposed an innovation on close-loop systems on CNC industrial machine based on process sensors exploiting supervised information of data acquired on the same machine and labelled by an expert operator.

This paper is presented as follows: in section 2 we give an overview of the present state of the art and the works that might have affinity to the current work; in section 3 we present the laser cutting procedure giving a brief overview of the technology to level unfamiliar readers; in section 4 we describe the proposed approaches, how we adapted the traditional networks to the current problem. The results of our extensive experiments are then reported in section 5 and finally we conclude with a discussion of our results in section 6 giving a peek to the possible future directions.

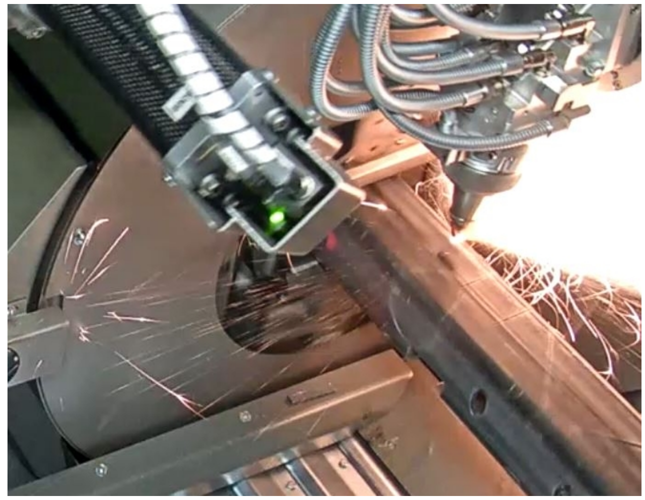
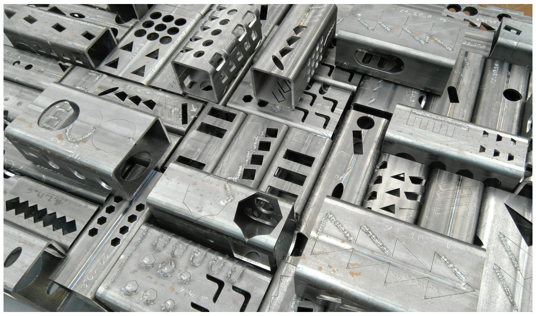


Figure 2. An example of visual outcome for a laser cutting process.

2. Related Works

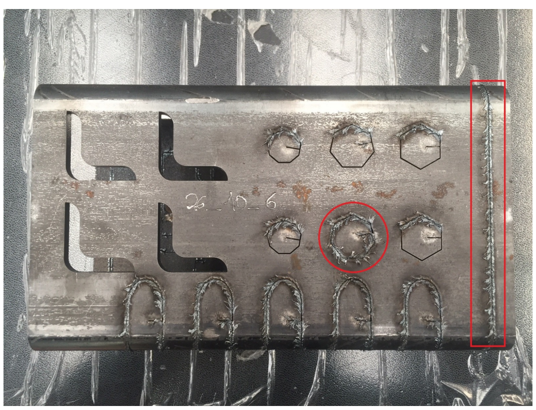
The advent of the fourth industrial revolution, also known as *Internet of Things* (a.k.a. commonly *IoT*) led to huge investments in the automation of industrial processes exploiting interconnected capabilities of different devices. Simple AI driven solutions to weakly automate industrial processes (e.g. thresholding, simple statistical analysis, etc.) have been adopted for a long time. Recently, advanced methods based on Machine Learning models are rapidly taking over [11] because despite their complexity they brought a large improvement in inference performance ensuring also high robustness to noisy input data and reducing critical failure to an acceptably low level [24].

Nowadays Machine learning finds wide application on a variety of fields such as healthcare applications [18, 21], security and tracking [25, 6], sport analysis [7, 4], societal challenges [9, 20]. The argument of this work is however different from the previously mentioned topics: in this case we propose an approach aiming at optimizing the industrial process, in order to monitor the ongoing quality of the production, rising alerts when the process is having setbacks with the opportunity to detect anomalies and possibly intervene in a way to restore the optimal process. In this regards, to the best of our knowledge this is the first work that uses Machine Learning in the field of the automation of the laser tube cut quality control. There are a few akin applications where Machine Learning is having a great impact, one of those is predictive maintenance. In this regard approaches of visual inspection [15] has been used to detect dirt in small conducts and cavities. In our case traditional imaging is not a good choice due to the high brightness of the laser beam while hitting the metal surface (as shown in



(a)

(b)



(c)

Figure 3. In (a) is shown a few examples of pieces and geometries cut to generate ALCIDE dataset, in (b) and (c) we show the result of an unexpected result, respectively an example of *Bad Cut* and a few *Missed Cut*. One can notice that in case of *Missed Cut* the geometry is not detached from the main piece.

Figure 2). We decided therefore to use machine parameters and laser reflectance detectors. Given the explorative flair of this research we have cast the problem as a classification problem between different possible outcomes using, as previously said, the information coming from the sensors installed on the machine. What we propose is not an uncommon strategy, similar decisions have been taken for problems regarding industrial predictive maintenance [24] unlike them in our case we only need to optimize one single aspect which is the quality of the cut.

3. Laser Cutting

The process of cutting a metal wall by melting a narrow line of material through the energy conveyed by a laser beam is called *laser cutting*.

At a process level, the laser beam — focused on a spot with diameter typically around 0.2 mm — melts the cut material and the liquid matter is quickly forced away from the cut by a high-pressure (4–20 bar) jet of gas, which solidifies the melt into little droplets. The cutting device is thus a *cutting head* that focuses the laser beam through a *nozzle*, which also conveys a pressurized gas flow coaxially to the laser beam. Depending on the cut material, the gas can be inert (mostly Nitrogen; for some specific application Argon) or oxidizing (typ. Oxygen). Inert gasses are used on materials that need to be protected from hot oxydization: in this case, pressure is typically high (15 bar or more) and also the laser power is relatively high. Oxygen is used for cutting materials that undergo exotermic oxydization (typically carbon steel): for these materials most of the melting power comes from the oxydization, the laser beam is more a driver of the melting process than its sole power source, and the laser power is typically low.

At a machine level, a laser cutting system is a Computer Numerical Control machine tool that uses a cutting head as a cutting tool. It is a machine that moves and possibly orient the cutting head in the 3-D space according to the trajectories written by the operator in a *part-program*. The machine architecture depends on the semi-finished material that has to be cut: when cutting sheet metal, the machine typically has two x–y orthogonal axes (on the sheet plane) plus a z axis (parallel to the beam) that moves to keep a constant distance between nozzle and sheet metal surface; when cutting tubes or beams the machine typically has at least three and up to 5 axes, in order to freely contour cuts on the 3-D surface of a — possibly not straight — tube.

The laser source — being it a delicate and heavy component — is typically placed in a separate, static box, which is connected to the cutting head via a suitable optical path. Nowadays the laser source and the optical path are both on optical glass fiber, while the previous generation of machines sported CO_2 laser sources and mirrored optical paths.

The machine used for the work here presented (Adige LT8.10, see Figure 1) is a 5-axes tube cutting machine, with 4 kW fiber laser source. The laser source has an optical combiner that allows to detect the backscattered light coming from the cut zone through the fiber and back to the laser source [19]. To the purpose of this work, we measure the light intensity by means of a photodetector and recorded it simultaneously to the other process parameters.

To serve as a guideline we propose an overview of the terminology used in the paper. We named *reading* each time measure of the set of parameters, for each reading is assigned a label according to the procedure described in section 3.1, a *sample* is a sequence of readings, a *geometry* is a single cut composed by a variable number of samples, such geometries are collected in a *piece* (a case full of pieces is shown in Figure 3 (a))

3.1. Adige Laser Cut Interruption DatasEt (ALCIDE)

Autonomous laser cut monitoring using process data may bring a boost in production performances and reliability of machines available on the market. However, being a niche application, there is no freely available data to test and compare our proposed approaches. For this reason we acquired a dataset using the laser cutting machine in Figure 1. Each reading is composed by the following set of parameters: the intensity of the light emitted by the cutting process, laser power, feed rate (i.e. cutting speed) and gas pressure set point. The dataset consists in 20 tube pieces (in Figure 3 (a) is shown part of the whole dataset), each of them composed by nearly 20 heterogeneous finite geometries, carried out on the flat surface and corners of the tube. The labelling has been done manually, by temporally locating areas where the tube presents plasma episodes (i.e. *Bad Cut*, see Figure 3 (b)) and where the cut is lost (i.e. *Missed Cut*) where the tube is welded (see Figure 3 (c)). The remaining class is considered *Good Cut* which is the optimal and expected outcome of the process.

The problem of creating such type of datasets is to define a strategy to generate samples of different classes that are actually representing cases that are close to real situations. The process to create bad examples is rather simple by using wrong machine presets (i.e. setting machine parameters for a thinner material) but this would include only a limited amount of cases and our target is to collect data that represents a large variety of cut interruptions. In order to do so we have produced negative examples by augmenting the cases of wrong preset selection mainly in two ways: by exploiting metal tubes with a pronounced seam that increases locally the thickness of the material and by using contaminated optics that introduce power losses and beam profile modifications, which in turn result in a sub-optimal optical configuration. We defined the possible classes as *Good Cut*, *Bad Cut* (i.e. when the cut is achieved but it is unstable, generally producing a low quality cut) and *Missed Cut* (i.e. when the cut kerf is closed and the laser does not pass through the material), we have seen these examples before in Figure 3.[1]

4. Proposed Approaches

In this section we outline the methods and the architectures we have adopted to classify the quality of the laser cut on metal tubes.

4.1. Gaussian Mixture Models

The first and simpler methodology is a generative approach based on Gaussian Mixture Models (GMM). We adopted this method because it is elegant, effective and widely known, it serves as a baseline comparison with more complex approaches. A generative approach ensures the possibility to sample from a probability distribution trained on a limited set of data. In our case the distribution is composed by a weighted set of components of K Gaussian densities as shown in Equation (1):

$$p(x) = \sum_{k=1}^{K} \pi_k \mathcal{N}(x|\mu_k, \sigma_k^2) \tag{1}$$

where as previously said, K is the number of components, $\mathcal{N}$ is the Gaussian component with μ and Σ as defining statistical parameters. For a more detailed description of the method see [2].

In our implementation we used the mean of each parameter in each sample and the standard deviation of the photodiode.

Most frequently the Gaussian mixture is used as an unsupervised learning algorithm to fit some unlabeled data. The degrees of freedom of the algorithm are the number of Gaussians that can be used to fit the model and the tolerance after which we consider the algorithm converged. In our case, we provide labels and therefore the algorithm is used in a supervised fashion.

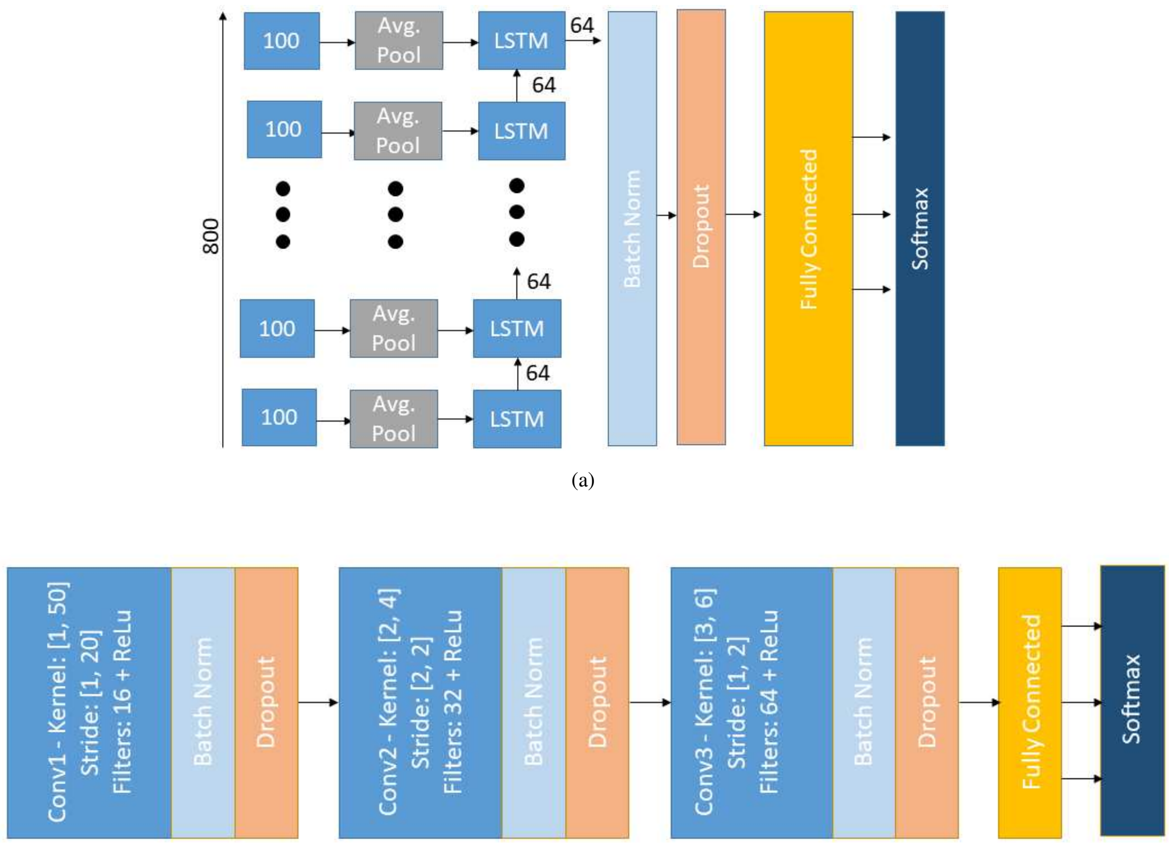
Data of each class are fit using the Expectation Maximization algorithm with a linear combination of five multivariate Gaussians. This number was a good trade-off between complexity and the quality of fitting. During the testing phase the trained Gaussian model is used to evaluate each sample for each class returning a set of posterior probabilities. The higher probability assigns the label to the sample.

4.2. Long Short Term Memory Network

The second approach is based on the Long Short-Term Memory Network [8], which is a particular type of Recurrent Neural Network where information is propagated along the sequence and trained to *remember* or *forget* the information at each time t, in each composing cell. This architecture is widely used for inferring from sequential data (e.g. skeleton tracking [13], saliency detection in videos [1], anomaly detection in industrial vision [15], etc.), while our case does not involve images but a set of parameters that are reflecting the status of the laser cutting machine and the outcome of the laser cutting procedure.

Our architecture is shown in Figure 4 (a). The sequence (i.e. *sample*) is fragmented in chunks of 100 subsequent

[1] ALCIDE dataset is available at https://research.promfacility.eu/#/dataset/alcide



(a)

(b)

Figure 4. The architectures of the proposed LSTM neural network (a) and the FCNN (b)

readings from which we extract parameter-wise the mean value using an average pooling layer. This preliminary operation is needed to stabilize the input parameters at time t. An experiment where each reading feeds a cell has been carried out producing very poor results. To our understanding we believe that having such a limited amount of fluctuating parameters constrains us to shorten the cells sequence of the RNN. However the need to cover a meaningful amount of time brought us to use average pooling on a branch of subsequent readings. The architecture that ensures better performances is composed by 8 LSTM cells that propagate a state vector of 64 elements. In cascade to the last LSTM cell we used batch normalization and dropout before connecting to a fully connected network where a Softmax activation function is applied to produce the soft-probability belonging to each class. The overall loss is computed using a traditional cross-entropy between labels and softmax logits.

4.3. Fully Convolutional Network

In the previous two methods we did not fully exploit the information of each reading, indeed the entire sequence is fragmented in chunks and the statistical parameters are extracted from each of them. In order to fully exploit the information of the single readings also in relation with other parameters at the same time, we propose an ad-hoc version of a Fully Convolutional Network (FCNN) [14] so that additional information regarding possible mutual information among parameters can emerge.

Our architecture is shown in Figure 4 (b). In the first layer we reduce the dimensionality of the sequence in a parameter-wise way, then in the second and third convolutional layers we use a kernel that convolves 2 and 3 different parameters together respectively. In both of these layers we further reduce the dimensionality of the sequence using temporal stride equal to 2. After *conv3* layer we placed a fully connected layer [2] that leads to the three classes probabilities. After every convolutional layer we use batch normalization and dropout to avoid overfitting effect.

In our ablation tests we tried different layer configurations without getting real noticeable improvements. The insertion of additional convolutional layers brings improve-

[2]The fully connected layer is implemented using a convolutional filter of the same size of the data at that layer, that's why we can still consider the architecture as a fully convolutional.

ment in training accuracy but introduces instability in the test phase denoting a clear overfitting effect.

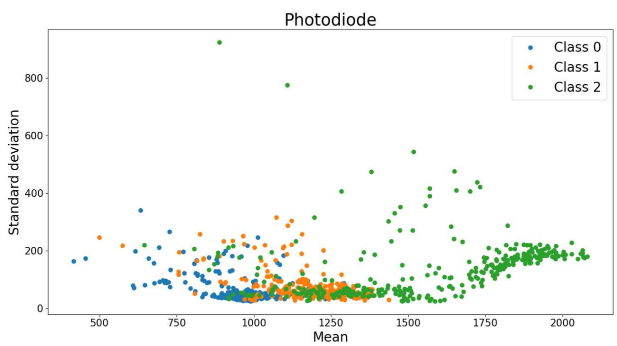
We have adopted the cross-entropy loss function as we did in the case of the LSTM approach.

5. Experimental Results

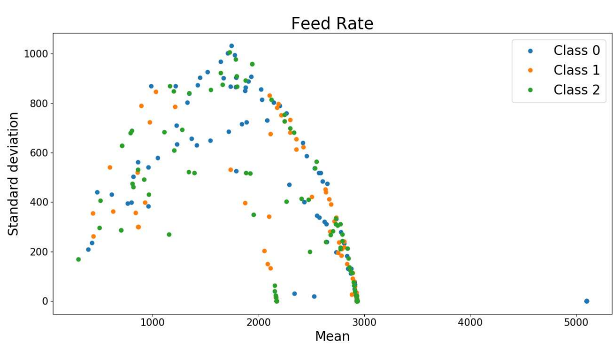
In this section we show the quantitative and qualitative results for the methods outlined in Section 4. Quantitative results are computed by balancing the dataset among the classes, splitting randomly 80/20 percent for training and test, reported results are the average result of 5 splits.

5.1. Weighted Max Voting

As long as we don't know much about of the acquired data, we first propose a simple classification method based on the statistics of each sample, aiming at analyzing the most significant parameters for the cut quality classification task. The methods simply reduces each sample to a tuple (μ, σ) for each parameter. For each parameter and for each class we then compute the average values of the tuple. In Figure 5 every sample is represented with a dot, different colors represent different classes.



(a)



(b)

Figure 5. The chart in (a) shows the dot plot for photodiode, each dot is a sample in the (μ, σ) space. Charts (b) reports the same plot for the feed-rate parameter. It is noticeable that in (a) classes are slightly separated, while in (b) they result are more confused and less informative. The classes *Good Cut*, *Bad Cut*, *Missed Cut* are encoded from 0 to 2 respectively.

From the visual inspection of the features, the photodiode signal (a) is the most informative and therefore we show this feature as reference in the qualitative images in Section 5.3. However, as shown in Figure 5 (a) this signal is not enough to clearly separate the classes. For example: *Good Cut* class, which is plotted in blue, is partially overlapped with the orange class that refers to the *Bad Cut*. The *Missed Cut* class is colored in green and it overlaps with the *Bad Cut* class.

These results are not surprising, the unstable process that produces a low quality cut most of the times results in a soft transition towards the loss of the cut. As further explanation it is true that the conceptual border between the classes is not as sharp and different annotators may label the readings differently.

In order to have a comparable result we used a small part of training data as validation set and we have computed the standalone accuracy for each parameter. We used this value as a weighting coefficient to be applied to the vote of each parameter as shown in Equation (2):

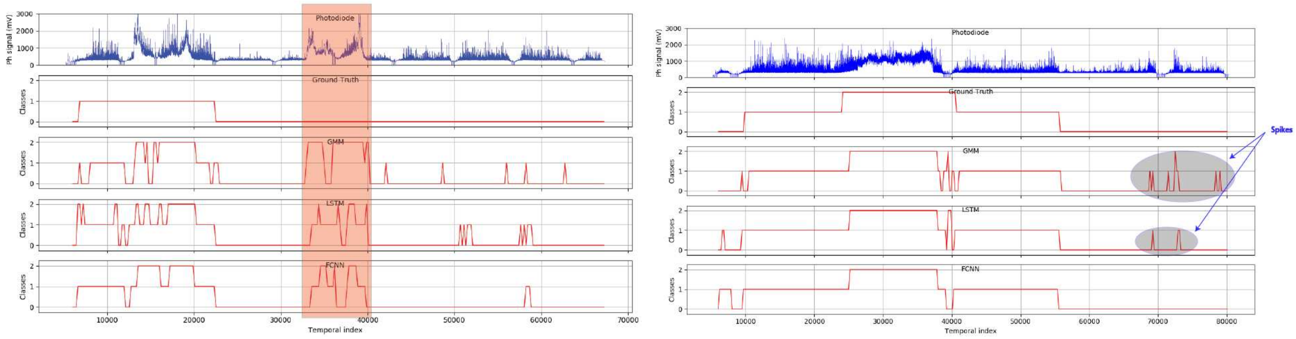
$$V_c = \sum_k \lambda_k v_k \tag{2}$$

where V_c is the voting score for class c as result of the summation over all the parameters of weighting coefficient λ and the unary function v for parameter k. Using this method we have reached 77% of classification accuracy improving the results of a simple voting algorithm of nearly 4%. It is noticeable that this method does not take into account the interdependence between parameters and the difference in contribution between them is pretty high and depends on the individual reliability.

5.2. Quantitative Results

In this section we show our model's results in an objective way using a standard evaluation procedure. The dataset has been balanced among classes and split by 80% training, and 10% each for validation and test. We did this procedure 5 times and reported the average accuracy. Different parameters have different unit of measurement, which is typical in a multimodal environment. In order to prevent small-values parameters to be overruled by other parameters with higher mean values we performed statistical standardization on each parameter. After this procedure our data resulted with mean 0 and standard deviation equal to 1 for all the parameters. Results are shown in Table 1.

Computational efficiency is also a key aspect for this type of applications. Table 2 shows the inference time for each method. All three methods have been tested on a regular laptop without GPU acceleration. The inference time is far below the time needed to acquire an entire sample



(a) (b)

Figure 6. Two examples of online prediction referred to the 3 proposed methods. The charts show the signal of the photodiode for an entire geometry with class ground truth and temporal inference. (a) highlights the problem of wrongly labelled examples where however the trained models are able to efficiently predict *Bad Cut*, while (b) shows the presence of spikes in the GMM and LSTM prediction but not in FCNN.

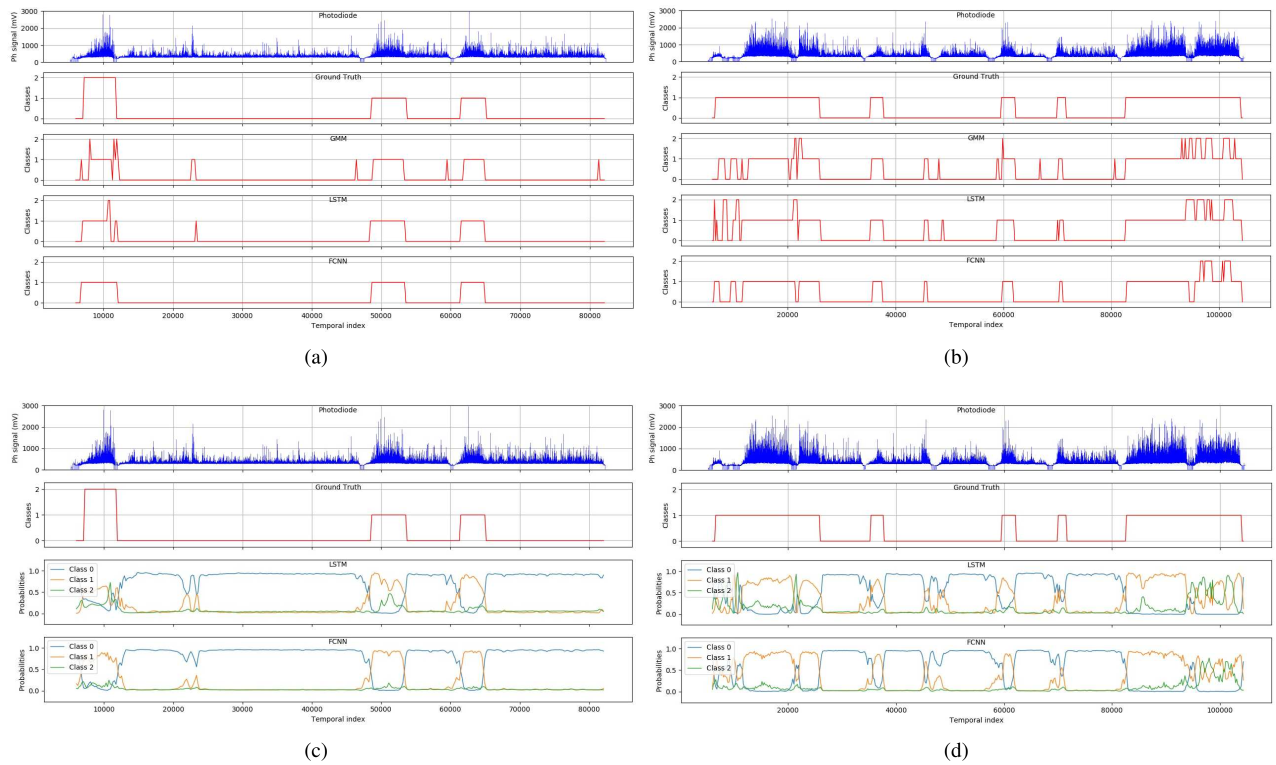


Figure 7. Here is shown two examples of online prediction referred to the 4 proposed methods. (c) and (d) better show the soft predictions obtained using the Neural Networks

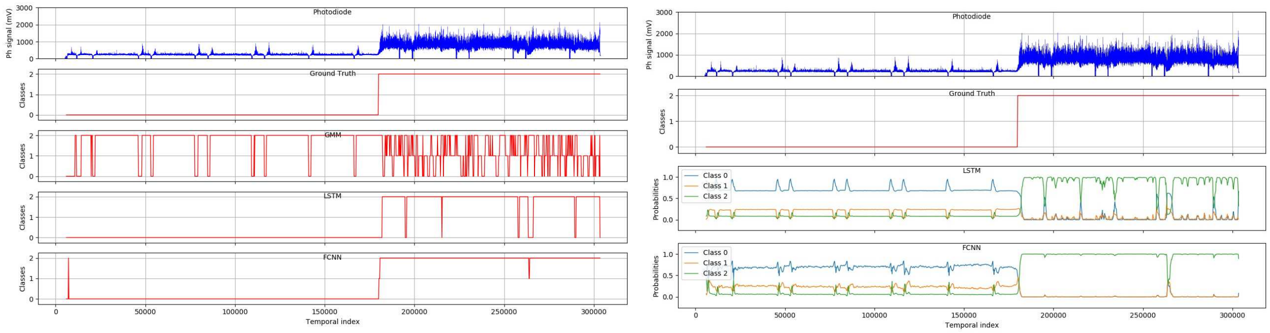
Table 1. Overall quantitative results in aggregate. Repeteability error refers to the standard deviation of 5 training/test splits.

Approach	Average Accuracy	Repeatability Standard Deviation
W-Voting	60.38	2.10
GMM	72.10	1.59
LSTM	76.15	1.40
FCNN	**76.65**	1.23

(i.e. 20 ms) and therefore directly deployable in the machine without further time optimization.

5.3. Qualitative Results

Quantitaive results may produce an objective reference for machine learning researchers, but when it comes to industrial process analysis a qualitative approach is often preferred. The qualitative analysis allows us to assess the machine learning model inference on a real laser cut, better



(a) (b)

Figure 8. An example of geometry acquired at different time with respect to the training dataset. Similarly to Figure 7 we show the hard predictions (a) and the probability transitions (b).

Table 2. Time inference for each sample (i.e. readings sequence) is reported in the table. Repeatability standard deviation for 5 repetitions is also reported for consistency assessment.

Approach	Time for sequence (ms)	Repeatability Standard Deviation
GMM	0.0042	0.0007
LSTM	1.2	0.3
FCNN	0.91	0.05

understanding where the model fails and what is necessary to do in order to improve the expected performance. Figures 6-7 shows a few examples of online predictions using the three proposed methods. For this experiment, unlike the quantitative case, we have split the dataset geometry-wise, using few complete geometries for the qualitative evaluation (i.e. entirely unseen by the training set). We used the trained models to produce an online estimation of the quality of the cut. Figure 6 reports the results of two different geometries. It is noticeable how Neural Networks based approaches are more stable comparing to the GMM one, in particular the FCNN presents much less spikes in the *Good Cut* class comparing to the other methods (see Figure 6 (b)). Figure 6 (a) shows an example of wrongly labelled sample: in this case all the approaches are able to detect the anomaly, which is supposed to be a *Bad Cut*. The higher stability of the FCNN model is made clearer from in Figure 7, where (a) and (b) show the hard results of another couple of geometries, while in (c) and (d) are shown the soft probabilities for the two geometries respectively. In the latter graphs we can see that uncertainty given by the FCNN is lower with respect to the one produced by the LSTM.

The last experiment addresses another issue that regards the periodic software updates released by the machine manufacturer for the laser cutting machines. These updates are released to improve the quality of the cutting presets and to introduce new functionalities. These modifications undermine the parameters used to train the models. In Figure 8 we show the results on a geometry cut performed four months after the dataset acquisition. The results show again how robust is a model trained using Neural Networks comparing to the GMM one which is performing way below the accuracy registered on the original dataset.

6. Conclusions

In this work we have shown an adaptation of well-known machine learning techniques applied to a novel application. In addition we are releasing a new dataset regarding laser-cut quality assessment procedure on industrial machines. We have shown three different approaches based on probabilistic, generative and discriminative models. From our experiments we have shown the higher quality and robustness of the approach based on Neural Networks, suggesting that even in different proportions, better results can be achieved by using a combination of different parameters in a convolutional settings, allowing better exploitation of intra-parameters relations. This observation also leads to a possible improved solution that uses a sort of residual architecture coupling all combinations of different parameters to fuse the information in a Neural Network bottleneck producing the inference.

We have shown the criticalities channeled by periodic software updates but so far we did not have the chance to test the trained models across different machines, on different productions (e.g. metal sheets) or using different technologies (i.e. oxidizing cut). These ideas will be developed in future works.

Acknowledgement

We would like to thank Adige S.p.A. for the opportunity of working on such interesting data and for their high enthusiasm and collaborative flair. The activity of Adige BLMGroup has been conducted within the framework of the research project LT4.0 according to the Regional Law LP6/99. This work has been also supported by Caritro Foundation.

References

[1] S. A. Bargal, A. Zunino, D. Kim, J. Zhang, V. Murino, and S. Sclaroff. Excitation backprop for rnns. In *CVPR*, pages 1440–1449. IEEE, 2018. 4

[2] C. M. Bishop. *Pattern recognition and machine learning*. springer, 2006. 4

[3] K. Chen, C. Zhang, C. Fang, Z. Wang, T. Bui, and R. Nevatia. Visually indicated sound generation by perceptually optimized classification. In *ECCV Workshops*, pages 560–574. Springer, 2018. 1

[4] S. Giancola, M. Amine, T. Dghaily, and B. Ghanem. Soccernet: A scalable dataset for action spotting in soccer videos. In *CVPR Workshops*, pages 1711–1721, 2018. 2

[5] R. Gomez, L. Gomez, J. Gibert, and D. Karatzas. Learning to learn from web data through deep semantic embeddings. In *ECCV Workshops*, pages 514–529. Springer, 2018. 1

[6] M. Guo, J. Lu, and J. Zhou. Dual-agent deep reinforcement learning for deformable face tracking. In *ECCV*, pages 768–783, 2018. 2

[7] M. Harmon, P. Lucey, and D. Klabjan. Predicting shot making in basketball using convolutional neural networks learnt from adversarial multiagent trajectories. *stat*, 1050:15, 2016. 2

[8] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997. 2, 4

[9] N. Jean, M. Burke, M. Xie, W. M. Davis, D. B. Lobell, and S. Ermon. Combining satellite imagery and machine learning to predict poverty. *Science*, 353(6301):790–794, 2016. 2

[10] V. V. Kniaz, V. A. Knyaz, J. Hladvka, W. G. Kropatsch, and V. Mizginov. Thermalgan: Multimodal color-to-thermal image translation for person re-identification in multispectral dataset. In *ECCV Workshops*, pages 606–624. Springer, 2018. 1

[11] P. Larrañaga, D. Atienza, J. Diaz-Rozo, A. Ogbechie, C. E. Puerto-Santana, and C. Bielza. *Industrial Applications of Machine Learning*. CRC Press, 2018. 2

[12] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *Nature*, 521(7553):436, 2015. 1

[13] J. Liu, A. Shahroudy, D. Xu, A. C. Kot, and G. Wang. Skeleton-based action recognition using spatio-temporal lstm network with trust gates. *T-PAMI*, 40(12):3007–3021, 2018. 4

[14] J. Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. In *CVPR*, pages 3431–3440, 2015. 2, 5

[15] S. Martelli, L. Mazzei, C. Canali, P. Guardiani, S. Giunta, A. Ghiazza, I. Mondino, F. Cannella, V. Murino, and A. Del Bue. Deep endoscope: Intelligent duct inspection for the avionic industry. *IEEE Transactions on Industrial Informatics*, 14(4):1701–1711, 2018. 1, 2, 4

[16] R. Montero, J. Victores, S. Martinez, A. Jardón, and C. Balaguer. Past, present and future of robotic tunnel inspection. *Automation in Construction*, 59:99–112, 2015. 1

[17] F. Piewak, P. Pinggera, M. Schäfer, D. Peter, B. Schwarz, N. Schneider, M. Enzweiler, D. Pfeiffer, and M. Zöllner. Boosting lidar-based semantic labeling by cross-modal training data generation. In *ECCV Workshops*, pages 497–513. Springer, 2018. 1

[18] M. Reiter, P. Rota, F. Kleber, M. Diem, S. Groeneveld-Krentz, and M. Dworzak. Clustering of cell populations in flow cytometry data using a combination of gaussian mixtures. *Pattern Recognition*, 60:1029–1040, 2016. 1, 2

[19] M. Sbetti, S. Bertoldi, D. Colombo, B. Previtali, G. Riva, M. Danesi, L. Molinari Tosatti, and D. Parazzoli. Method for controlling a laser cutting process and laser cutting system implementing the same, 2012. European Patent Office Number: WO2012143899. 2, 3

[20] S. Sokolov-Mladenović, M. Milovančević, I. Mladenović, and M. Alizamir. Economic growth forecasting by artificial neural network with extreme learning machine based on trade, import and export parameters. *Computers in Human Behavior*, 65:43–45, 2016. 2

[21] A. J. Stangl, E. Kothari, S. D. Jain, T. Yeh, K. Grauman, and D. Gurari. Browsewithme: An online clothes shopping assistant for people with visual impairments. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 107–118. ACM, 2018. 1, 2

[22] M. Teichmann, M. Weber, M. Zoellner, R. Cipolla, and R. Urtasun. Multinet: Real-time joint semantic reasoning for autonomous driving. In *2018 IEEE Intelligent Vehicles Symposium*, pages 1013–1020. IEEE, 2018. 1

[23] G. Williams, P. Drews, B. Goldfain, J. M. Rehg, and E. A. Theodorou. Information-theoretic model predictive control: Theory and applications to autonomous driving. *IEEE Transactions on Robotics*, 34(6):1603–1622, 2018. 1

[24] T. Wuest, D. Weimer, C. Irgens, and K.-D. Thoben. Machine learning in manufacturing: advantages, challenges, and applications. *Production & Manufacturing Research*, 4(1):23–45, 2016. 2, 3

[25] C. Zhang, H. Li, X. Wang, and X. Yang. Cross-scene crowd counting via deep convolutional neural networks. In *CVPR*, pages 833–841, 2015. 2