

# Extending electronic shot counters with machine learning-based ammunition discrimination for targeted maintenance

Journal of Defense  
Analytics and  
Logistics

Nathan Morsa

*Department of Electrical Engineering and Computer Science, University of Liège,  
Liège, Belgium and  
FN e-novation, FN Herstal, Herstal, Belgium*

Received 13 February 2025  
Revised 16 April 2025  
25 April 2025  
Accepted 2 May 2025

## Abstract

**Purpose** – Weapon fleets degrade differently depending on usage. Electronic shot counters provide armorers with accurate tracking and allow more efficient preventive maintenance practices. We propose a machine learning technique for general-purpose ammunition type discrimination, thereby enhancing targeted maintenance.

**Design/methodology/approach** – We study an experimental shot counter deployment to understand its impact on maintenance practices. We then extend the existing EDGAR machine learning technique to solve the discrimination problem in a generic way on a weakly-labelled dataset, requiring only the total counts for each shot type. By repurposing intermediate neural network activations, we simplify training and minimize computational overhead. We evaluate our approach against a widely used live/blank discrimination algorithm.

**Findings** – We show a 94% improvement in instance-level error rate and perfect burst-firing performance. This generalizes across weapon platforms without hyperparameter adjustment. The feature incurs as little as an 8% overhead (2.8 ms on a 64 MHz ARM Cortex-M4F).

**Originality/value** – Compared to existing techniques, it promises applicability to a broader range of weapon configuration discrimination tasks, including platforms previously deemed too complex or constrained. Additionally, we examine how performance scales with dataset size to offer practical data collection guidelines, a major challenge in this field. This technique supports a new generation of shot counters and targeted maintenance, thereby reducing costs, preventing incidents and increasing operational availability.

**Keywords** Electronic shot counters, Gunshot detection, Ammunition discrimination, Preventive maintenance, Operational availability, Weakly supervised learning

**Paper type** Research paper

## 1. Introduction

The acquisition and ongoing maintenance of small arms, along with the procurement of spare parts and accessories, represent substantial investments for modern armies. The management and upkeep of this weapon fleet demands considerable time and directly impacts the reliability of weapons, a factor critical to mission success and the safety of the user. Paradoxically, small arms management is one of the rare areas of military logistics that has not yet entered the digital era. Current tracking relies on manual paper logs and spreadsheets, leading to frequent errors that compromise maintenance schedules.

One solution to this problem is the systematic use of electronic shot counters. These devices continuously and accurately track weapon usage and maintenance status.

© Nathan Morsa. Published in the *Journal of Defense Analytics and Logistics*. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) license. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this license may be seen at <http://creativecommons.org/licences/by/4.0/legalcode>

This work was supported by FN Herstal, which provided funding, weapons, dedicated time for the research and facilitated data collection. FN Herstal had no influence on study design, data analysis, interpretation of the results or the decision to publish.



Journal of Defense Analytics and  
Logistics  
Emerald Publishing Limited  
e-ISSN: 2399-6447  
p-ISSN: 2399-6439  
DOI 10.1108/JDAL-02-2025-0004

Military applications require resistance to harsh environmental conditions. This includes full submersion in water, and direct exposure to dust, grease, saline environments and various types of chemicals. For NATO-associated countries, required testing is described in the STANAG 4370 standard. They should also be fully integrated to prevent accidental or intentional removal, and require no specific maintenance and in particular no battery replacement for a prolonged period of time.

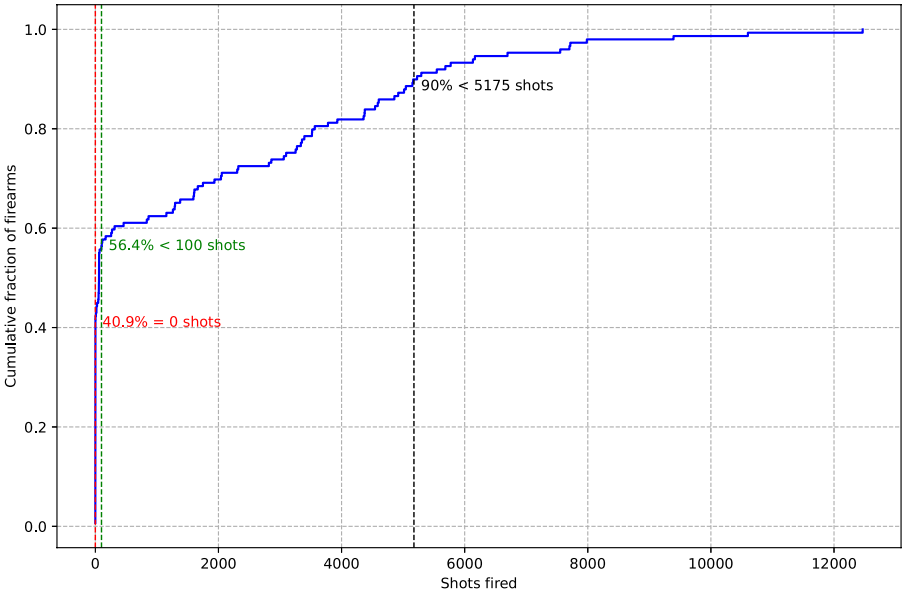
A possible solution is the use of magnetic sensors attached to the moving parts of the weapon and monitoring their displacement. Examples include the FN SmartCore for FN SCAR by [FN Herstal \(2025\)](#) described in a patent by [Gering \(2020\)](#), the Heckler and Koch Bolt Motion Sensor System ([Roth, 2024](#)) described in a patent by [Staiger et al. \(2023\)](#), and the Radetec RISCpro ([Radetec, 2024](#)). While they provide high integration and the possibility of battery-free operation, they generally require significant modifications of the firearm to be installed. Among these, only the FN SmartCore for FN SCAR allows discriminating between live and blank ammunition, although through the use of an additional accelerometer.

Most popular solutions rely on micro-electromechanical systems (MEMS) accelerometers and are permanently moulded into the weapon handle or securely fastened to it. Examples include the FN SmartCore for FN Minimi and FN M3M by [FN Herstal \(2025\)](#), WeaponLogic devices by [Senseer \(2024\)](#), and the PAL by [LMT Advanced Technologies \(2022\)](#). On top of shot counting, all of these provide capabilities to discriminate the ammunition type between live and blank rounds, with the PAL also detecting the use of a suppressor.

1.1 Operational impact of electronic shot counters

To evaluate the cost savings that electronic shot counters can provide, 150 machine guns of a modern army unit were equipped with electronic shot counters for a year. The shot counting results are reported in [Figure 1](#).

We can observe that over 40% of the weapons had not been fired at all, and over 56% were barely used. This very uneven usage across the fleet presents an opportunity for optimization.



**Figure 1.** Distribution of the number of shots fired in real-world usage. Usage of 150 machine guns was tracked over a full year

Significant time savings can be obtained by clearing minimally operated weapons without inspection, speeding up inventorying operations, and facilitating maintenance.

Furthermore, frequent tracking of weapon usage through shot counters allows armourers to plan ahead maintenance operations and spare parts ordering, reducing associated costs. By also reducing maintenance downtime, the operational availability of the fleet increases, allowing fewer weapons to be procured while still ensuring sufficient availability.

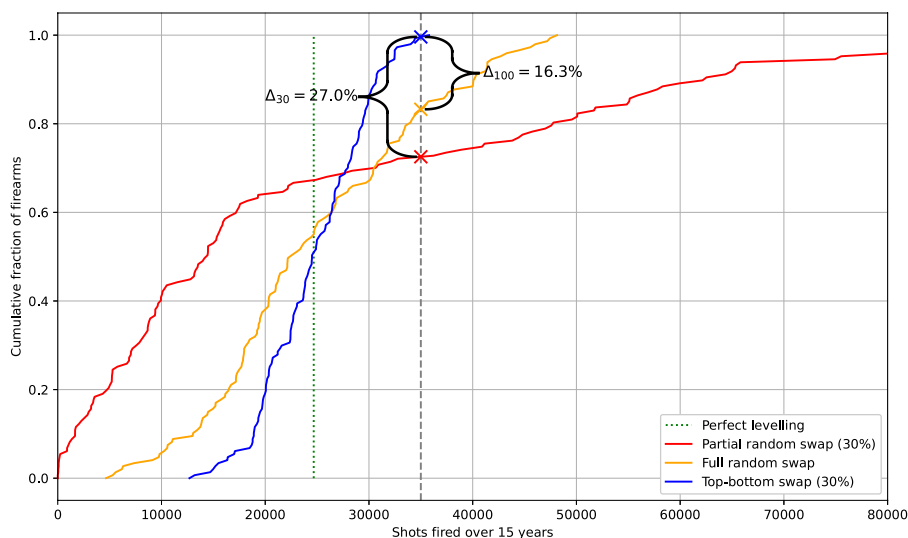
Tracking the number of shots also allows for improved wear-levelling as illustrated in Figure 2. Assuming that a firearm would fail after an arbitrary 35,000 rounds, a strategy informed by shot counters (“top-bottom swap”) where each year the top 15% in usage of the fleet is swapped with the bottom 15% shows a 27% difference in attrition after 15 years over a random strategy with the same number of swaps. This is equivalent to a gain of 3 years in fleet attrition level.

Finally, access to consolidated statistics allows fleet managers to leverage lessons learned from the field and learn the actual profile of parts consumption depending on the type of training or the field operations performed. With wide adoption, live tracking of the ammunition stock is also possible. Furthermore, usage data and in particular the schedule of firing allow fleet managers to evaluate if the mandated doctrine is actually practised in the field, and whether it is still appropriate.

### 1.2 Ammunition discrimination for targeted maintenance

Once a reliable shot counting is achieved, more subtle signal variations can be leveraged to perform shot discrimination, which is the detection of the type of ammunition fired, or the weapon configuration in which it was fired. This enables more targeted maintenance.

In particular, the use of blank rounds (training ammunition containing only powder and no projectile) causes incomplete powder combustion leading to increased deposits of residues inside the mechanism compared to standard live rounds. The projectile of live rounds, however, creates permanent copper deposits inside the barrel, eventually leading to a loss of accuracy. This also heats up the barrel significantly, leading to thermal fatigue. While the mobile parts are solicited similarly with both types of ammunition, live ammunition causes increased wear on the locking mechanism compared to blank ammunition. In short, blanks



**Figure 2.** Wear-levelling strategies comparison. Data from Figure 1 is extrapolated over 15 years, and end of life is studied at 35,000 shots

induce less long-term damage on some parts of the weapon but increase the build-up of residue, requiring more frequent clean-ups.

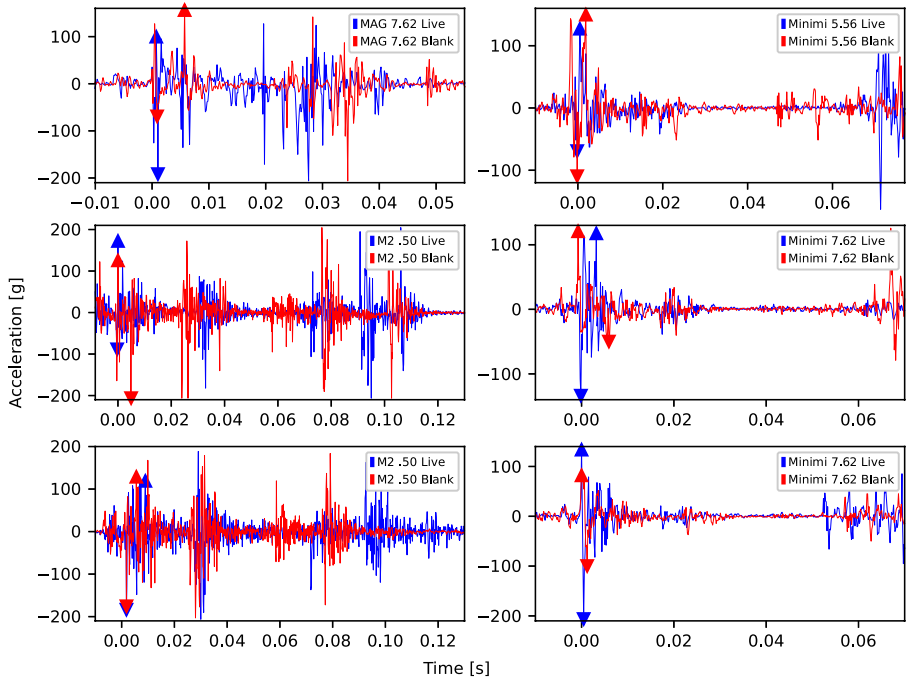
By detecting and individually tracking these different usages through electronic shot counters, the type and location of the degradation can be further accurately estimated. Recommendations for a more effective maintenance procedure can therefore be provided.

## 2. Discrimination problem and state of the art

The exact acceleration curve resulting from a shot depends on a large number of unknown external variables (ammunition powder load, parts lubrication and temperature, use of a weapon mount or support, shooter stance and grip tightness, mounted accessories weight, etc.), making it challenging to isolate a single metric for ammunition discrimination. A few examples are given in Figure 3.

Due to the specific nature of the problem of shot counting, as well as the customary level of trade secrecy in this industry, publicly available research on shot counting remains limited. A significant number of important patents exist on the topic, particularly for accelerometer-based methods (Johnson and Kulesza, 2005; Joannes *et al.*, 2010; Ufer and Brinkley, 2014; Weiss *et al.*, 2021; Boettcher *et al.*, 2024; Morsa, 2023a; Asbach and Canty, 2024). Since ammunition discrimination relies on initial shot detection, studies in this area are even scarcer, and public information is mostly available through patents.

A patent by Staiger *et al.* (2023) states that “[ammunition discrimination] can be executed, for example, on the basis of the different acceleration pulses of moving parts such as the breech



**Figure 3.** Example inputs (high-resolution images available in the digital version). Pairs of live/blank rounds were acquired under identical, controlled conditions. Each pair is aligned on an arbitrary significant peak in the signal (at  $t = 0$ ).  $P_P$  and  $P_N$  are denoted by  $\blacktriangle$  and  $\blacktriangledown$ , respectively. The distinction between live and blank ammunition is obscured across different weapon systems due to substantial differences in firing cycles. Even within a single weapon, minor configuration adjustments or the intrinsic variability between individual rounds can lead to significant variations in behaviour

or the different recoil pulses on the overall system”, although a method to do so is not described.

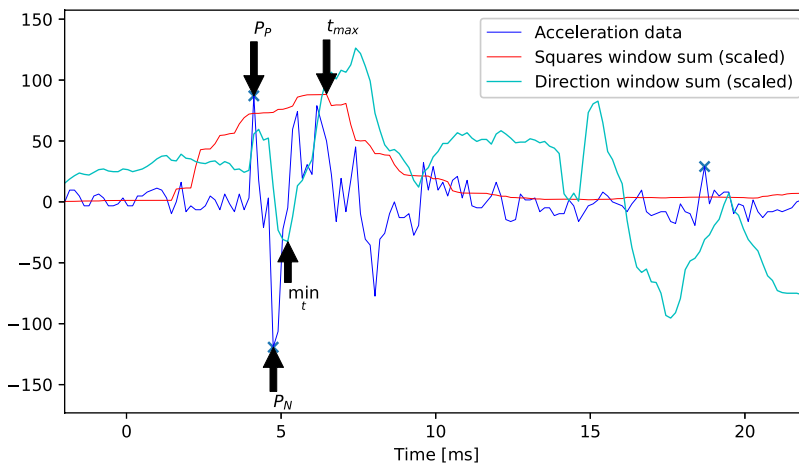
A patent by Joannes *et al.* (2010) describes an approach for discrimination between live and blank rounds, in which the “direction of the first initial shock” is used. A related patent by Gering (2020) goes into further details, describing a blank round as one “having a phase of acceleration to the front for several milliseconds prior to the recoil.” In this technique, an acceleration signal is considered as a blank round when “the recoil acceleration comprises two components of opposite directions.” This difference is created by the way the ejection of the projectile mass (or lack thereof) affects the recoil of the weapon. However, due to the large vibrations created during the ejection and the difficulty of acquiring and processing an acceleration signal of sufficiently high resolution, detecting the direction of the initial shock is challenging. In particular, always-on sensors that rely on a low-power system (such as a piezoelectric element or a low-power accelerometer) to wake up from deep sleep are usually not able to capture the first milliseconds of a shooting event which, as just described, contain most of the information about the ammunition type used.

The aforementioned patents implement this by first detecting that a significant acceleration event is happening. Then, following the convention that accelerations are positive toward the front of the weapon, looking for the largest acceleration values in the positive ( $P_P$ ) and negative ( $P_N$ ) directions, as illustrated in Figure 4.

The direction of the initial shock according to this metric is then computed by looking at the sign of the difference between these peaks, optionally offset with a parameter  $P_0$  as such:  $\Delta P = P_0 + (P_P - P_N)$ . Although intuitive, this method does not always yield the desired results. As an example, this approach fails on the Minimi 5.56 Live example in Figure 3 where the positive peak is significantly higher than the negative one, resulting in a “Blank” misclassification by this technique.

An alternative method for computing the initial shock direction which has shown strong discrimination properties is to look for the maximum recoil before the point of highest acceleration intensity ( $t_{\max}$ ). This can be computed with the following, where  $a[t]$  is the input acceleration time series,  $t_0$  the start of the current acceleration event, with  $D_0$ ,  $w_1$ , and  $w_2$  as tunable parameters:

$$t_{\max} = \arg \max_t \sum_{i=-w_1/2}^{w_1/2} (a[t + i])^2, \quad (1a)$$



**Figure 4.** Example of classical ammunition discrimination. Metrics are computed from the raw acceleration data, and from moving windows on this signal

$$D = D_0 + \min_{t \in [t_0, t_{\max}]} \sum_{i=-w_2/2}^{w_2/2} a[t + i]. \quad (1b)$$

Discrimination is then performed by examining the sign of  $D$ , with  $D_0$  serving as an adjustable decision threshold.

By considering the absolute values of  $\Delta P$  and  $D$  as confidence measures for each technique’s discrimination capability, they can be combined into an ensemble metric,  $C$ , as follows, where  $\alpha$  is a tunable parameter:

$$C = \begin{cases} \Delta P, & \text{if } \alpha(\Delta P)^2 > D^2, \\ D, & \text{otherwise.} \end{cases}$$

Discrimination is then finally performed by examining the sign of  $C$ . Decisions are first performed individually, then then overall class is assigned based on the majority result across instances. While this approach provides very high accuracy for sufficiently long streaks of shots, its performance often declines over short, isolated bursts, typically dropping to around 75% accuracy (see [Section 8](#)). Whether this is significant depends heavily on the weapon type.

Over recent years, machine learning methods have proven increasingly effective at generating state-of-the-art algorithms for this type of classification problem. [Welsh and Roy \(2017\)](#) are the first to apply it to gunshot activity detection. They then associate with [Khan et al. \(2018\)](#) for another proof-of-concept in firearm type detection. In 2021, [Welsh et al. \(2021\)](#) introduce a distance feature based on Dynamic Time Warping, which enables a generic comparison of samples with a template without the need for feature selection by a domain expert. Their method achieves perfect firearm recognition and moderate success in discriminating between three types of live ammunition per weapon. However, their small dataset and fixed experimental setup limit generalisability, as changes in weapon configuration can fundamentally alter the shape of the acceleration curve, undermining template-based approaches. Additionally, the technique is computationally intensive and unsuitable for real-time use on ultra-low-power devices.

[Loeffler \(2014\)](#) identified parallels with the domain of fall detection in human activity. Later work by [Putra et al. \(2017\)](#) studies accelerometer-based fall detection by decomposing the impact event into sub-events in the time domain, thereby showing similarities with shot counting, and applying machine learning techniques. [Santos et al. \(2019\)](#) apply deep-learning techniques to this same problem, and propose an approach based on a convolutional neural network suitable for embedded use.

In 2023, Morsa published a paper ([Morsa, 2023b](#)) and an associated patent ([Morsa, 2023a](#)), detailing the motivation for using machine learning techniques to produce shot-counting algorithms. The work introduces a method for training compact networks suitable for use in embedded devices from datasets requiring minimal labelling effort, allowing the acquisition and use of datasets multiple orders of magnitude larger than those used in previous studies. In this publication, we aim to extend this technique with ammunition discrimination capabilities, while maintaining the same philosophy of minimal labelling effort and low overhead appropriate for real-time inference in constrained embedded environments.

### 3. Evaluating discrimination performance

#### 3.1 Available datasets

Training data was acquired on the FN Minimi 5.56, FN Minimi 7.62, FN MAG, and FN M2HB-QCB, in an identical manner to the original EDGAR proposal. For the Minimi 5.56, data was collected using both short and long barrels at various gas return settings, including extreme values. Most measurements were taken with mid-life barrels that were used until decommissioning, with additional data gathered from new barrels. The weapon was fired

from the shoulder, hip, bipod, and tripod, both without accessories and with added accessories totalling 3 kg. The Minimi 7.62 was tested using a new barrel across multiple gas return settings, including extremes. It was fired from the shoulder, hip, and bipod, with and without 3 kg of accessories. For the FN MAG, data was gathered using three weapons, one new and two nearing decommissioning, alternated across trials. Each was set to different gas return levels, including extreme settings. Firing configurations included the shoulder, bipod, tripod, elastic mount, and two fixed mounts. Accessories totalling 3 kg were used in all configurations except the shoulder. The M2HB-QCB was tested in live fire using one new and one nearly-decommissioned barrel, employed alternately. For blank firing, a specialised blank-firing barrel was required due to the constraints of the short-recoil system. The weapon was fired from a tripod, fixed mount, elastic mount, and a deFNder teleoperated station. Feeding was done using both a minimal 10-round belt and a standard 100-round belt to account for the additional mass.

The acceleration curves in our datasets typically consist of 15 shots for the Minimi and MAG variants and 10 shots for M2. This ensures that the counting error rate can be reliably reported. Each weapon was tested across four firing sequences: isolated shots, small random bursts of 3–6 rounds, bursts of four shots spaced by single shots, and continuous single bursts. Only live or blank rounds are fired in a single acquisition, allowing for a full evaluation of the discrimination error.

The available data used for training and testing are reported in [Table 1](#). Two examples of devices used for their acquisition are shown in [Figure 5](#). As discussed in more depth in the

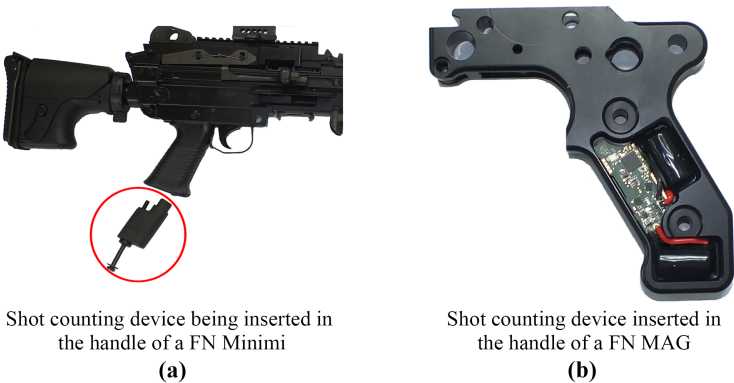
**Table 1.** Summary of the number of shots per dataset

	Minimi 5.56 Learning	Validation	Test	Minimi 7.62 Learning	Validation	
Live	4,596	1,920	1,800	4,707	943	
Blank	2,370	839	1,150	1,800	348	
Non-shot	>400	>55	>0	>97	>35	

	MAG Learning	Validation	Test	M2 Learning	Validation	Test
Live	9,400	2,156	1,755	4,923	669	260
Blank	4,609	1,335	990	3,839	499	250
Non-shot	>1,067	>206	>192	>1,261	>174	>30

**Source(s):** The author



**Figure 5.** Illustration of the shot counting devices used in this study

original publication (Morsa, 2023b), accurate weak labels cannot be provided for the number of non-shot events. Instead, only a lower bound on their count is reported, based on deliberately conducted events such as dry firings. Since these events naturally occur frequently during firing, their actual number is significantly higher.

### 3.2 Methodology

Since we know the actual type of each shot in our dataset, we can define a balanced discrimination error rate  $\bar{E}_D$  as:

$$\bar{E}_D = \frac{1}{N} \sum_i E_{D_i}, E_{D_i} = \frac{\sum_{j \neq i} \mathbf{d}_{ij}}{\sum_j \mathbf{d}_{ij}} \quad (2)$$

where  $\mathbf{d}_{ij}$  represents the count of shots (or other countable events, if applicable) of type  $i$  that are classified as type  $j$  by the model. For the simplified case of discrimination between live and training ammunition, this becomes:

$$\bar{E}_D = \frac{E_B + E_K}{2}, E_B = \frac{\mathbf{d}_{BK}}{\mathbf{d}_{BB} + \mathbf{d}_{BK}}, E_K = \frac{\mathbf{d}_{KB}}{\mathbf{d}_{KK} + \mathbf{d}_{KB}} \quad (3)$$

where B (for **B**all) denotes live rounds and K (for blan**K**) denotes blank rounds.

This definition ensures that our error metric is bounded within the interval  $[0, 1]$ , and minimizes instances in which the discrimination error is artificially increased or decreased due to counting errors. Since we do not know the real-world distribution of the different types of ammunition, averaging across types prevents artificially improving the discrimination error rate by adding data focused on well-performing cases. Similarly, we adapt the counting error from Morsa (2023b) into a version balanced across ammunition types, defined as:

$$\bar{E}_C = \frac{1}{N} \sum_i E_{C_i}, E_{C_i} = \frac{\sum_{x \in \mathcal{A}_i} |\hat{\mathbf{c}}_x - \mathbf{c}_x|}{\sum_{x \in \mathcal{A}_i} \mathbf{c}_x} \quad (4)$$

where  $\mathcal{A}_i$  denotes the set of acceleration time series for ammunition type  $i$ ,  $\mathbf{c}_x$  is the real number of shots in time series  $x$ , and  $\hat{\mathbf{c}}_x$  is the predicted number.

## 4. Learning from weak labels

Automatic firearms generate numerous unrelated acceleration events during firing, their number and characteristics varying significantly across weapon configurations. This makes accurate counting in a wide range of setups a challenging task. Manually annotating each shot with strong labels requires firearms expertise, is time-consuming, and remains susceptible to human error. The original proposal for shot counting (Morsa, 2023b) introduced a method based on learning from label proportions, enabling an instance-level classifier to learn from weak labels in a dataset labeled only with the number of shots in a time series. These aggregate labels can be obtained inexpensively and reliably by counting expended rounds, especially when using pre-numbered ammunition boxes or magazines. This approach greatly simplifies data acquisition, a major challenge of this field, and enables the practical collection of datasets orders of magnitude larger. As will be shown later, the discrimination task benefits significantly from larger datasets making the EDGAR technique a strong foundation for extension with ammunition discrimination capabilities.

Given a time series containing a shooting sequence, where  $\mathbf{c}_i$  represents the number of occurrences of each event category and  $\mathbb{X}$  denotes the division into candidate slices, we can



define the proportions  $\mathbf{p}_i = \frac{c_i}{|\mathbb{X}|}$ . The key is to reframe the counting problem as a category proportion problem, formulating part of the loss function  $\mathcal{L}_{prop}$  as a comparison between the actual proportions and those derived from the predictions of the machine-learning-based classifier  $\hat{\mathbf{p}}$ . While the original results are restricted to two categories, *non-shot* and *shot* ( $N = 2$ ), the method presented here accepts an arbitrary number of categories. In this paper, we study an example of ammunition discrimination that classifies between *non-shot*, *shot with live ammunition* (or *live*), and *shot with training ammunition* (or *blank*) leading to three categories ( $N = 3$ ). The introduced loss function and corresponding neural network can be easily extended from counting to discrimination by increasing the number of classes ( $N$ ) considered.

However, as the *non-shot* class becomes one of many, the loss function increasingly emphasises the discrimination accuracy over the counting accuracy, which is counterproductive for the general shot counting problem. Indeed, information on weapon usage and degradation from accurate shot detection is far more important than that provided by discrimination. To address this issue, we introduce a new hyperparameter  $\alpha$ , to balance these competing objectives:

$$\mathcal{L}_{prop} = \mathcal{L}_{counting} + \alpha \mathcal{L}_{discrimination}, \alpha \in ]0, 1]. \quad (5)$$

Assuming that  $\mathbf{p}_1$  represents the proportion of *non-shot* candidates, we define new sets of proportions to recover the original counting problem by summing the proportions of all *shot-type* classes:

$$\mathbf{p}' = \left( \mathbf{p}_1, \sum_{i=2}^N \mathbf{p}_i \right), \hat{\mathbf{p}}' = \left( \hat{\mathbf{p}}_1, \sum_{i=2}^N \hat{\mathbf{p}}_i \right). \quad (6)$$

The counting loss is then defined using these reduced proportions:

$$\mathcal{L}_{counting} = - \sum_{i=1}^2 \mathbf{p}'_i \log(\hat{\mathbf{p}}'_i) + \sum_{i=1}^2 \mathbf{p}'_i \log(\mathbf{p}'_i). \quad (7)$$

The discrimination loss follows the original formulation:

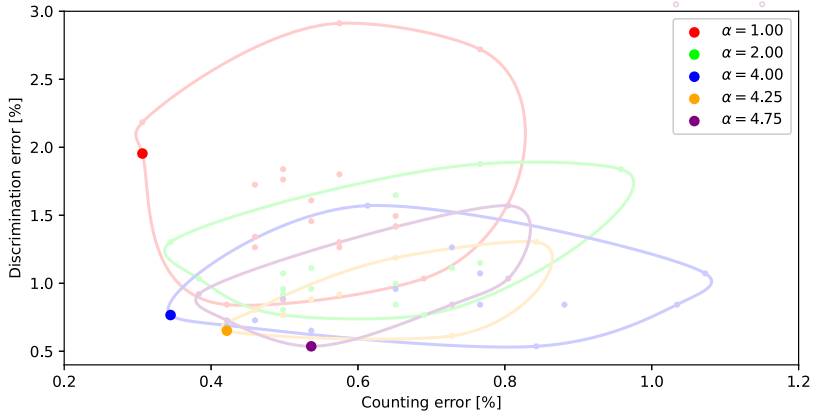
$$\mathcal{L}_{discrimination} = - \sum_{i=1}^N \mathbf{p}_i \log(\hat{\mathbf{p}}_i) + \sum_{i=1}^N \mathbf{p}_i \log(\mathbf{p}_i). \quad (8)$$

The value of  $\alpha$  can then be adjusted, revealing a Pareto front between counting and discrimination accuracy, from which an optimal balance can be selected, as illustrated in [Figure 6](#).

## 5. Separated networks

Dividing the network into a two-step process, as illustrated in [Figure 7](#), yields significant improvements.

In this new structure, the first part of the network is tasked solely with the counting sub-problem. Its structure is identical to the one described by [Morsa \(2023b\)](#), which is itself an adaption of the one by [Santos et al. \(2019\)](#). The hyperparameters were either kept identical or tuned using the procedure advised in the methodology section of the proposal. If the candidate is confirmed as a shot, the input vector is passed to a second network tasked with the discrimination sub-problem. This network mirrors the architecture of the first, except for the final layer, where the number of neurons corresponds to the number of discrimination classes. This design enables a direct comparison with the monolithic network in terms of classification capabilities. This technique offers several advantages:



**Figure 6.** Error rates on an unseen test dataset for clouds of 20 randomly initialized runs with different values of  $\alpha$ . The contour of each cloud is outlined and non-dominated solutions are highlighted

### 5.1 Computational budget

Separating the two sub-problems clarifies the distribution of computational budget between the counting and discrimination sub-problems. The budget for the second network is derived from the remaining computational power remaining after allocation to the higher-priority counting network. Moreover, the discrimination network only needs to run on candidates confirmed as shots by the counting network, thus reducing the total average inference time. If *shot* candidates represent a proportion  $\lambda$  of the total number of candidates, then, for a fixed maximum average inference time  $\bar{T}_{\max}$  and a counting inference time  $\bar{T}_C$ , the maximum average discrimination inference time  $\bar{T}_D$  can be estimated as follows:

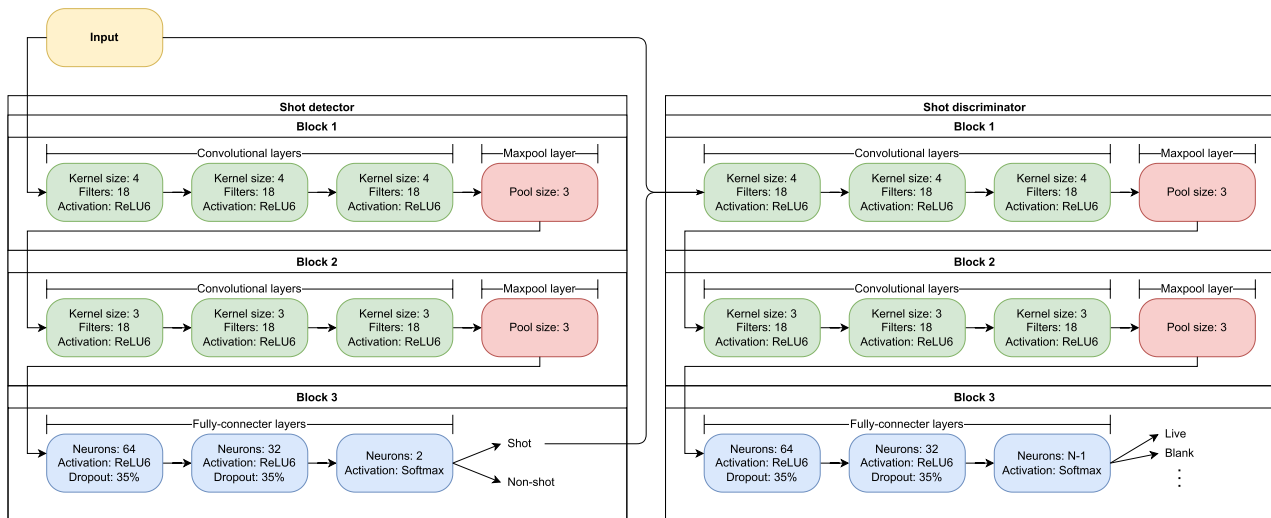
$$\bar{T}_{\max} \geq \lambda(\bar{T}_C + \bar{T}_D) + (1 - \lambda)\bar{T}_C \Leftrightarrow \bar{T}_D \leq \frac{\bar{T}_{\max} - \bar{T}_C}{\lambda}. \quad (9)$$

A good estimate of  $\lambda$  will therefore allow for a larger  $\bar{T}_D$  budget. A more precise estimate of the discrimination budget can be obtained through individual inspection of the time series from a representative input sample for their density in candidates, and their *shot/non-shot* proportion. It is then possible to determine what percentage of the input population can be satisfied with a given discrimination budget. An iterative optimization can then be applied to find the maximum discrimination budget that still meets the minimum required percentage of the input series.

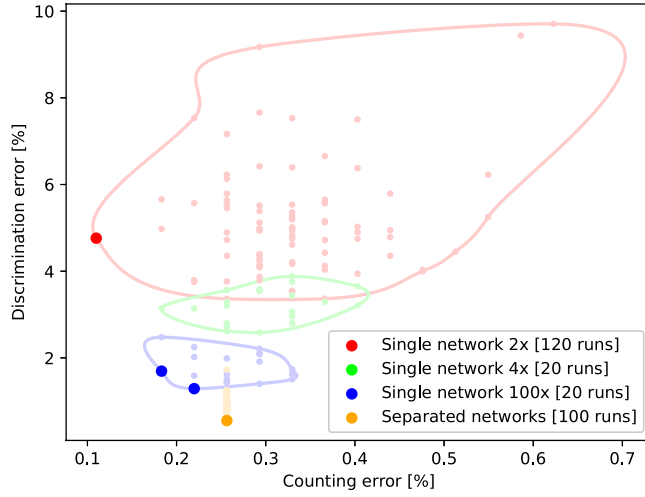
### 5.2 Discrimination performance

[Figure 8](#) compares the proposed two-step process with large single networks. Larger networks are obtained by proportionally increasing the number of convolution channels in the convolutional layers and the number of neurons in the dense layers. In the case of separated networks, the counting network is derived first, yielding a fixed counting error for all associated discrimination networks. The most direct comparison is with a single network that has the same total number of parameters as the separated networks, which is twice the number of parameters in the original counting network. This is represented in red in [Figure 8](#). Importantly, while these models are identical in size, the larger single model must be run in full and requires twice the inference time for *non-shot* candidates.

We observe that the separated discrimination network performs significantly better, achieving a reduction in the median discrimination error of more than 80%. However, we also



**Figure 7.** Two-step network structure for discrimination. The exact number of parameters depends on the available computational budget



**Figure 8.** Error rates on an unseen test dataset for clouds of randomly initialized runs, comparing large single networks against a two-step process. The contour of each cloud is outlined and non-dominated solutions are highlighted

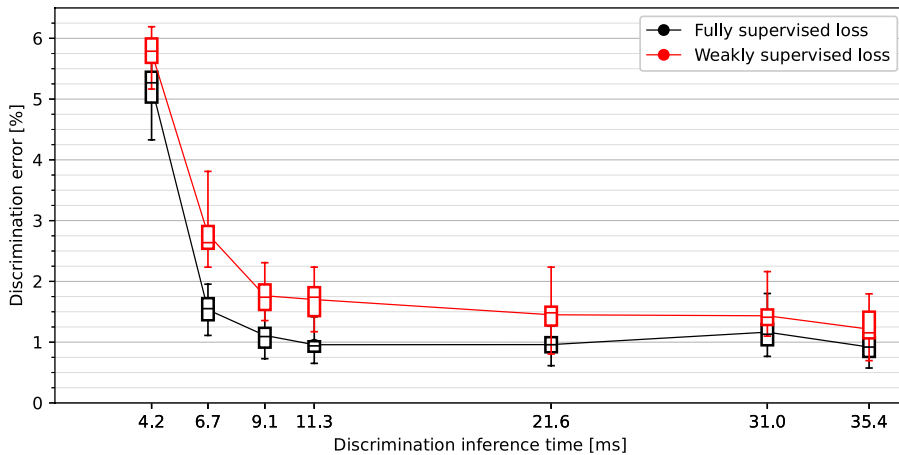
observe that a very small proportion of networks show a slight decrease in counting error. These outliers have likely allocated the increased network size towards the counting sub-problem rather than the discrimination sub-problem. As expected, increasing the size of the single network improves performance in both counting and discrimination. However, even networks with 100 times more parameters still exhibit a median discrimination error 86% higher compared with the separated networks. These results suggest that isolating the two sub-problems enhances the effectiveness of the discrimination part significantly.

### 5.3 Ease of training

Separating the two sub-problems allows their respective networks to be defined and trained independently. The higher-priority counting model can thus be derived first and locked to ensure optimal accuracy, independent of discrimination performance. To reduce development lead time, a key reason for the use of machine learning in shot counters, one could consider deploying a counting-only model before deriving the discrimination model. This additional feature could then be deployed later as a firmware update.

In addition, as illustrated in Figure 8, attempts to reduce the discrimination error by increasing the network capacity might instead result only in improved counting performance, necessitating further adjustment of  $\alpha$  to ensure that the increased capacity is allocated to the discrimination sub-problem. Separating the networks ensures that any supplementary capacity given to the discrimination network will be used for that purpose.

The original proposal has the primary advantage of being able to be trained using only weak labels on time series that may contain mixed classes, thus reducing the need for extensive labelling efforts. However, when constructing a dataset under controlled conditions, it is possible to ensure that each time series contains only a single type of ammunition. Since the counting network separates the hard-to-label *non-shots*, the remaining *shot* vectors can be labelled inexpensively on an individual basis (though this will also include the false positives given by the counting network). The problem addressed by the discrimination network then becomes fully supervised, enabling us to leverage a more usual formulation of the loss function, such as cross-entropy. The main benefit is a reduction of the training time by approximately 65%. The median discrimination error is also reduced by 20%, as shown in the rightmost boxes of Figure 9. Notably,



**Figure 9.** Discrimination error rates on an unseen test dataset for batches of 20 randomly initialized networks at various levels of  $D$ . The highest inference time corresponds to fully separated networks. Performance between fully and weakly supervised losses is also compared. Inference time is evaluated on a 64 MHz Cortex-M4F microcontroller

this represents a difference of only 0.23% in absolute discrimination error, which remains stable across different discrimination network capacities as obtained by the technique described in [Section 6](#) with higher error rates. This suggests that, generally, using weakly supervised learning with the EDGAR technique does not significantly penalise prediction performance.

## 6. Shared activations

While the previous section has shown that splitting the two sub-problems leads to several benefits, we can assume that some redundancy will exist between the two networks, as they both operate on the same input vector to solve a similar problem. Since parallelisation of the two sub-problems is often unavailable on resource-constrained embedded hardware, we can reduce the inference time and network size of the discrimination network by repurposing intermediate activations of the counting network.

The initial layers of the convolutional neural network, prior to the first pooling operation, are the most computationally expensive, accounting for about 60–80% of the inference time in typical use cases.

Depending on the specific problem and available resources, two hyperparameters are introduced:

- (1) The depth at which the counting network is reused ( $D$ ), measured in layers from the input. Deeper copies further reduce inference time and memory footprint, but also reduce prediction performance as discrimination-specific information is not recovered from the original signal.
- (2) Optionally, layers before  $D$  can be replicated in a narrower fashion (fewer convolution channels/neurons) before being concatenated to the copied activations. These extra channels/neurons, whose number is denoted as  $E$ , allow the final network to recover some of the discrimination-specific information that may be missing from the activations of the counting network.

An example of the resulting structure is given in [Figure 10](#). Maintaining identical structures for the two subnetworks ensures that variations in  $D$  and  $E$  result in a controlled comparison between shared and non-shared layers.

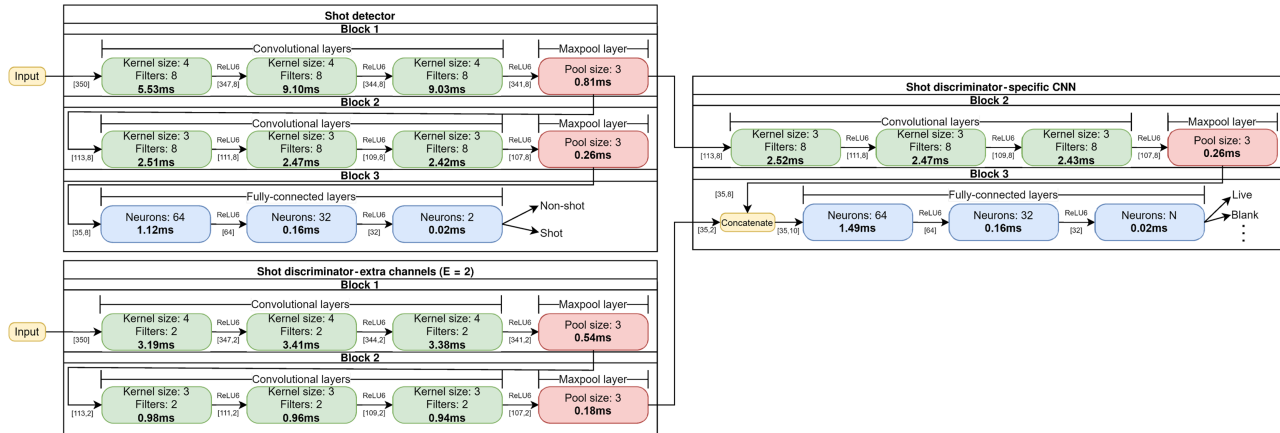


Figure 10. Example of a network with shared activations used for MAG discrimination with  $D = 4$  and  $E = 2$

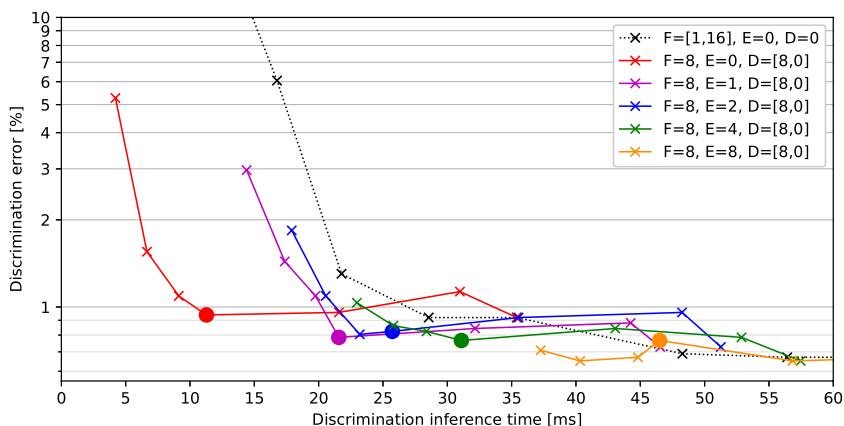
Training is accomplished by building the whole network structure, then copying and freezing the weights of the layers shared with the counting network. At inference time, the counting sub-graph is traversed first up to the counting outputs. If a *shot* is predicted, the discrimination sub-graph, starting from the shared activations (and optionally the input of the extra layers), is traversed separately.

In Figure 11, we study how the discrimination error rate and inference time evolve when  $D$  and  $E$  are tweaked. As a basis, we compare this with how the error/time tradeoff is otherwise performed by varying the number of convolution channels, denoted as  $F$ . For reference, the associated counting network runs in 33.9 ms. Looking first at variations of  $D$  for each fixed  $E$ , we observe that increasing  $D$  (from right to left) up to  $D = 4$  has only a minor impact on discrimination performance, with an average increase of 12% and as low as 2% for  $E = 0$ . Meanwhile, inference time is reduced by 25 ms, representing a 68% decrease in the  $D = 0$  case. After that point, further increases of  $D$  generally increase the error rate exponentially while providing relatively lower inference time gains.

Comparing with the reference approach of managing the error/time tradeoff by varying the number of convolution channels  $F$  (shown in dotted black), we notice that with  $E = 0$ , much faster networks are obtained for a similar error rate. Considering a maximum acceptable error rate of 5%, the shared activations technique allows us to run in 4.2 ms instead of 16.8 ms, or a 75% reduction. Considering that shots represent about 67% of the evaluated candidates in this dataset, the ammunition discrimination feature can be added with as little as 2.8 ms per candidate, representing an 8% overhead on the counting inference time. At a desirable 1% error rate, discrimination runs in 11.3 ms instead of 28.5 ms, representing a 60% reduction.

Looking now at the different curves obtained when  $E$  is increased (from left to right), we observe an expected reduction in discrimination error at the cost of increased inference time. Comparing with the reference, for a restricted computational budget of around 35 ms similar to that of the counting network, we notice that the discrimination error can be reduced by up to 20%. For unrestricted inference time, the difference becomes progressively smaller as the “extra” network becomes sufficient to capture the full discrimination sub-problem independently.

Based on these results, our general recommendation is to use shared activations with  $D = 4$  and  $E = 0$  as a starting point. From there, either  $E$  can be increased until the maximum acceptable inference time is reached for best prediction performance, or  $D$  can be increased until the maximal acceptable error is reached for best inference time performance.



**Figure 11.** Logarithmic discrimination error rates and inference time on an unseen MAG test dataset at various levels of  $D$ ,  $E$ , and  $F$ . Points at  $D = 4$  are highlighted. Each point is the median of 20 runs.  $D = 3$  and  $D = 7$  are identical to  $D = 4$  and  $D = 8$  respectively. Inference time is evaluated on a 64 MHz Cortex-M4F microcontroller

The main drawback of this approach is that the discrimination network is directly tied to a specific counting network. Any change to the latter will likely require retraining the former. We thus recommend keeping fully separated networks in the case of less constrained environments.

## 7. Weapon configuration majority voting

There is a physical limitation on how quickly a weapon can be switched from one configuration to another, particularly when specific accessories need to be mounted on the platform. Notably, switching from firing live rounds to blank rounds typically requires attaching a blank-firing adaptor to the end of the barrel for both functional and safety reasons. A similar operation is required when switching to and from the use of a suppressor. We can therefore assume that all shots detected in a burst (i.e. uninterrupted automatic firing) must belong to the same weapon configuration.

In addition, affixing the required accessory requires the operator to wait until the barrel cools down to safe handling temperatures. Therefore, provided that all intervals between detected shots are shorter than this minimum configuration switching time, the detected shots should all belong to the same weapon configuration.

This information can be leveraged to improve the overall discrimination performance by performing a majority vote on the detected shot classes. Individual predictions are accumulated until there is a period without detected shots that exceeds the minimum time required to switch configurations. At that point, all pending predictions can be permanently recorded with the ammunition type that corresponds to the majority of the predictions.

## 8. Performance

As a performance comparison baseline, the technique described in [Section 2](#) has been implemented. This is the best publicly available classical implementation and has been deployed in multiple commercially successful products as part of the maintenance workflow of several armies worldwide. Its parameters  $P_0$ ,  $D_0$ ,  $\alpha$ ,  $w_1$ , and  $w_2$  have been manually tuned on each of the learning and validation datasets to provide best possible performance. Classical discrimination algorithms are paired with their classical counting algorithms, EDGAR discrimination models with EDGAR counting models. The only exception is the M2 case, where the EDGAR discrimination model is paired with the classical counting algorithm. Classical discrimination algorithms all use majority voting to improve their final performance. In practical use, this would typically be performed on a time basis; however, majority voting is here only applied within each acceleration curve, which represents a more restricted usage. The results are presented with and without this technique, enabling separate observation of its impact and offering insight into single-shot discrimination performance. The results are reported in [Tables 2 to 6](#). The balanced counting error, averaged across ammunition types, is reported as  $\bar{E}_C$ . The number of detections on non-shot inputs is reported as  $FP$ . The Minimi 5.56 test set was acquired independently of the training data and, as a result, shows a distribution shift, as it includes some configurations which were not captured in the training data. Therefore, results on the validation set are also reported. The Minimi 7.62 does not have a test set, as its independent acquisition has not yet been completed; thus, results on the validation set are reported instead. Considering that the datasets largely over-represent extreme cases, classical algorithms show acceptable performance for practical usage of discrimination on the Minimi platforms, particularly in 7.62. The MAG and M2 cases, however, exhibit much higher error rates, making their use unreliable. In contrast, EDGAR discrimination achieves perfect performance when using majority voting. Additional insight can be gained by examining individual predictions, which remain important for sporadic usage. The Minimi 5.56 shows the highest reported individual error rate, with  $\bar{E}_D$  at 3.74%, due to the inclusion of configurations not represented in the training data. However, this remains



<i>Classical (raw)</i>		$d_{iB}$	$d_{iK}$	$\bar{E}_D$	25.52%
$\bar{E}_C$	1.32%	$d_{Bj}$	448	$E_B$	23.74%
$FP$	30/>55	$d_{Kj}$	228	$E_K$	27.31%
<i>Classical (voted)</i>		$d_{iB}$	$d_{iK}$	$\bar{E}_D$	17.06%
$\bar{E}_C$	1.32%	$d_{Bj}$	203	$E_B$	10.76%
$FP$	30/>55	$d_{Kj}$	195	$E_K$	23.35%
<i>EDGAR (raw)</i>		$d_{iB}$	$d_{iK}$	$\bar{E}_D$	0.60%
$\bar{E}_C$	0.03%	$d_{Bj}$	7	$E_B$	0.36%
$FP$	0/>55	$d_{Kj}$	7	$E_K$	0.83%
<i>EDGAR (voted)</i>		$d_{iB}$	$d_{iK}$	$\bar{E}_D$	0.00%
$\bar{E}_C$	0.03%	$d_{Bj}$	0	$E_B$	0.00%
$FP$	0/>55	$d_{Kj}$	839	$E_K$	0.00%
<b>Source(s):</b> The author					

**Table 3.** Benchmark data: Minimi 5.56 – test set

<i>Classical (raw)</i>			$d_{iB}$	$d_{iK}$	$\bar{E}_D$	33.55%
$\bar{E}_C$	1.57%	$d_{Bj}$	1,279	479	$E_B$	27.25%
$FP$	0/>0	$d_{Kj}$	459	693	$E_K$	39.84%
<i>Classical (voted)</i>			$d_{iB}$	$d_{iK}$	$\bar{E}_D$	10.81%
$\bar{E}_C$	1.57%	$d_{Bj}$	1,758	0	$E_B$	0.00%
$FP$	0/>0	$d_{Kj}$	249	903	$E_K$	21.61%
<i>EDGAR (raw)</i>			$d_{iB}$	$d_{iK}$	$\bar{E}_D$	3.74%
$\bar{E}_C$	1.38%	$d_{Bj}$	1,702	55	$E_B$	3.13%
$FP$	0/>0	$d_{Kj}$	50	1,097	$E_K$	4.36%
<i>EDGAR (voted)</i>			$d_{iB}$	$d_{iK}$	$\bar{E}_D$	0.00%
$\bar{E}_C$	1.38%	$d_{Bj}$	1,757	0	$E_B$	0.00%
$FP$	0/>0	$d_{Kj}$	0	1,147	$E_K$	0.00%
<b>Source(s):</b> The author						

**Table 4.** Benchmark data: Minimi 7.62 – validation set

<i>Classical (raw)</i>		$d_{iB}$	$d_{iK}$	$\bar{\mathbf{E}}_{\mathbf{D}}$	8.89%
$\bar{E}_C$	1.23%	$d_{Bj}$	58	$E_B$	6.25%
$FP$	0/>35	$d_{Kj}$	40	$E_K$	11.53%
<i>Classical (voted)</i>		$d_{iB}$	$d_{iK}$	$\bar{\mathbf{E}}_{\mathbf{D}}$	3.23%
$\bar{E}_C$	1.23%	$d_{Bj}$	60	$E_B$	6.47%
$FP$	0/>35	$d_{Kj}$	347	$E_K$	0.00%
<i>EDGAR (raw)</i>		$d_{iB}$	$d_{iK}$	$\bar{\mathbf{E}}_{\mathbf{D}}$	0.64%
$\bar{E}_C$	0.14%	$d_{Bj}$	4	$E_B$	0.42%
$FP$	0/>35	$d_{Kj}$	346	$E_K$	0.86%
<i>EDGAR (voted)</i>		$d_{iB}$	$d_{iK}$	$\bar{\mathbf{E}}_{\mathbf{D}}$	0.00%
$\bar{E}_C$	0.14%	$d_{Bj}$	0	$E_B$	0.00%
$FP$	0/>35	$d_{Kj}$	349	$E_K$	0.00%
<b>Source(s):</b> The author					

**Table 5.** Benchmark data: MAG – test set

<i>Classical (raw)</i>			$d_{iB}$	$d_{iK}$	$\bar{E}_D$	40.13%
$\bar{E}_C$	3.86%	$d_{Bj}$	1,245	380	$E_B$	23.38%
$FP$	3/>192	$d_{Kj}$	562	426	$E_K$	56.88%
<i>Classical (voted)</i>			$d_{iB}$	$d_{iK}$	$\bar{E}_D$	38.48%
$\bar{E}_C$	3.86%	$d_{Bj}$	1,261	364	$E_B$	22.40%
$FP$	3/>192	$d_{Kj}$	539	449	$E_K$	54.55%
<i>EDGAR (raw)</i>			$d_{iB}$	$d_{iK}$	$\bar{E}_D$	0.73%
$\bar{E}_C$	0.56%	$d_{Bj}$	1,739	6	$E_B$	0.34%
$FP$	2/>192	$d_{Kj}$	11	978	$E_K$	1.11%
<i>EDGAR (voted)</i>			$d_{iB}$	$d_{iK}$	$\bar{E}_D$	0.00%
$\bar{E}_C$	0.56%	$d_{Bj}$	1,745	0	$E_B$	0.00%
$FP$	2/>192	$d_{Kj}$	0	989	$E_K$	0.00%
<b>Source(s):</b> The author						

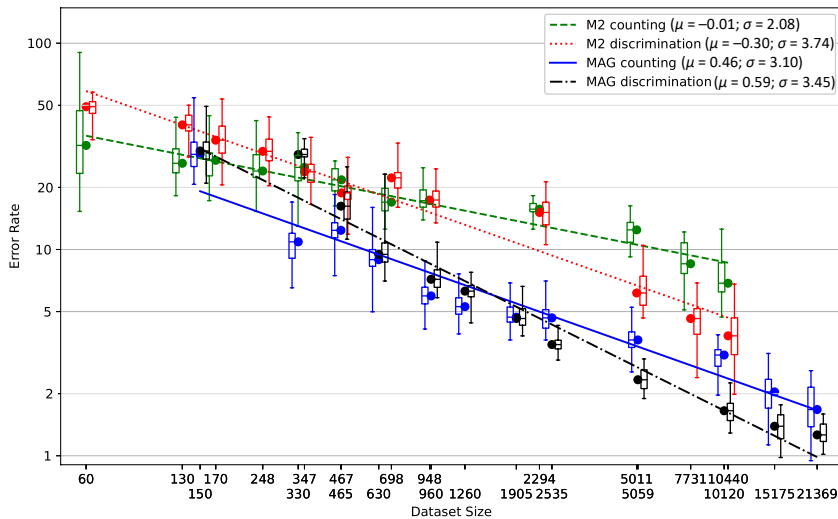
**Table 6.** Benchmark data: M2 – test set

<i>Classical (raw)</i>			$d_{iB}$	$d_{iK}$	$\bar{E}_D$	45.05%
$\bar{E}_C$	1.80%	$d_{Bj}$	98	162	$E_B$	62.31%
$FP$	0/>30	$d_{Kj}$	67	174	$E_K$	27.80%
<i>Classical (voted)</i>			$d_{iB}$	$d_{iK}$	$\bar{E}_D$	54.74%
$\bar{E}_C$	1.80%	$d_{Bj}$	80	180	$E_B$	69.23%
$FP$	0/>30	$d_{Kj}$	97	144	$E_K$	40.25%
<i>EDGAR (raw)</i>			$d_{iB}$	$d_{iK}$	$\bar{E}_D$	1.58%
$\bar{E}_C$	1.80%	$d_{Bj}$	259	1	$E_B$	0.38%
$FP$	0/>30	$d_{Kj}$	5	175	$E_K$	2.78%
<i>EDGAR (voted)</i>			$d_{iB}$	$d_{iK}$	$\bar{E}_D$	0.00%
$\bar{E}_C$	1.80%	$d_{Bj}$	260	0	$E_B$	0.00%
$FP$	0/>30	$d_{Kj}$	0	180	$E_K$	0.00%
<b>Source(s):</b> The author						

well below the 10% error rate commonly accepted for practical discrimination purposes, and represents a 65% improvement over the classical error rate with majority voting, and 89% without. A similar improvement of 93% is observed in the validation set for the 7.62 variant. Even greater improvements are observed in the MAG and M2 use cases, with respective improvement of 98 and 96% and error rates falling around 1% before voting. On average, the EDGAR technique shows a 94% reduction in individual error rate. This not only makes deployment of the ammunition discrimination feature possible but also promises a high degree of reliability, even in the case of isolated shots.

## 9. Minimal required dataset

Neural network performance relative to dataset size has been successfully modelled by a power law across a wide range of domains and scales (Hestness *et al.*, 2017; Kaplan *et al.*, 2020; Henighan *et al.*, 2020). In Figure 12, we can observe that this relationship holds true for both the counting and the discrimination sub-problems, in both the MAG and M2 datasets. For this study, majority voting is disabled, as the voted error rate otherwise quickly falls to zero. The discrimination behaviour starts off with a worse error rate for very small datasets, but is more responsive to the increase in dataset size. Interestingly, the MAG dataset for the discrimination sub-problem shows a deviation from a power law at low levels of



**Figure 12.** Error rate versus dataset size with power-law model fitting on the median values (log-log scale). Hyperparameters are fixed over the whole range. Reduced datasets are drawn from the full dataset with balanced representation across weapon configurations, setups, and ammunition types, ensuring the preservation of overall diversity

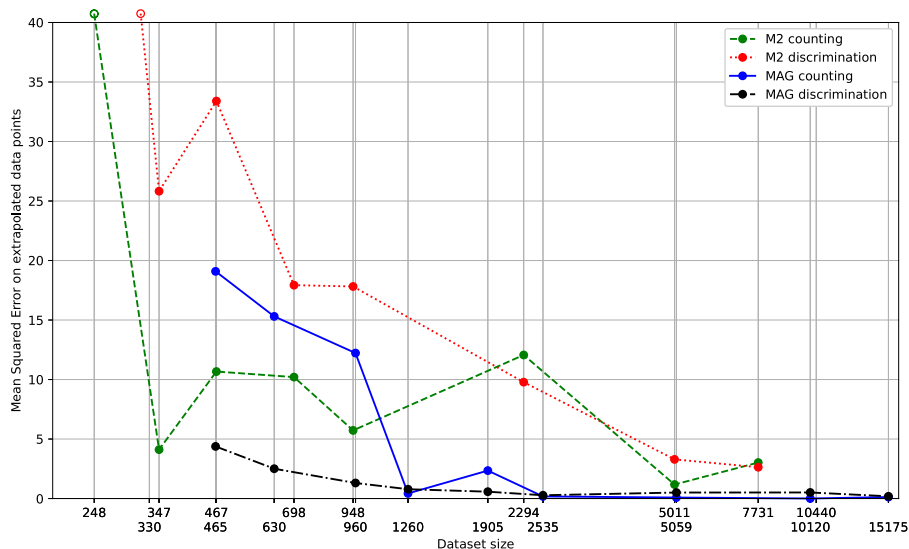
discrimination error. This deviation is expected from the known literature as we approach the saturation of our model (due to its fixed size) around  $\bar{E}_D = 1\%$ .

This information has a very practical use for our application. Each new weapon requires a new dataset, and acquiring the samples is expensive, time-consuming, and often requires significant planning to ensure the availability of hardware and testing grounds. Quantifying in advance the necessary number of shots to achieve adequate performance is therefore of great interest. [Rosenfeld et al. \(2019\)](#) apply extrapolation to the power law deduced from smaller datasets to deduce the performance given by a bigger dataset on a fixed benchmark.

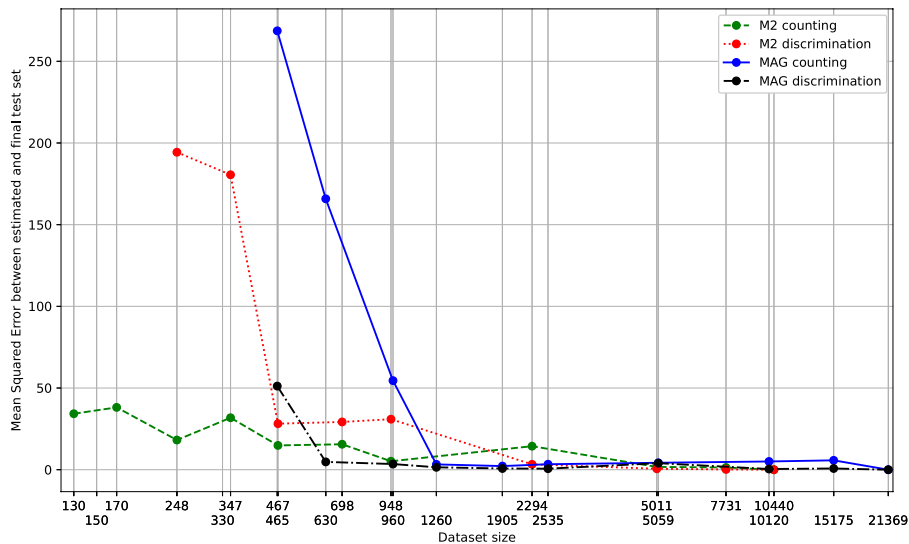
In our case, an independent benchmark is generally unavailable during preliminary testing of a new firearm and must therefore be constructed alongside our learning datasets. In [Figure 13](#), we observe the error rate between the extrapolated values and the actual values as estimated on the largest available dataset. We generally observe a sharp drop before a more gradual improvement. [Figure 14](#) gives us further insight into this behaviour by comparing the error rates obtained on the test sets of the reduced datasets and those obtained on the test set of the biggest available dataset. This represents the error in the estimation of the error rates on which the decay model will be fitted. Before a sufficiently low error is reached, the reduced test dataset is not yet representative of the full distribution and significantly underestimates the actual error rate, rendering any fit, and thus extrapolation, inaccurate.

From [Figure 14](#), we can thus make a general recommendation for a dataset creation procedure that includes preliminary testing with 1,000–2,000 shots in diversified setups, allowing the test set to reach representativeness. We hypothesise that the exact value required is proportional to the number of possible configurations of the weapon. This is supported by the MAG counting dataset being the last to reach representativeness, as this firearm offers a very high number of possible configurations and, thus, high variance in possible behaviours under use.

Using this exploratory dataset, one can fit the power law and estimate the number of shots required to achieve the desired counting and discrimination performance. Examples are provided in [Table 7](#), where the smallest dataset that reaches representativeness is used to provide predictions for a desired error rate  $E$ . We observe that, generally, the extrapolated



**Figure 13.** Mean-square error between extrapolated values and test values against dataset size. Extrapolated values are obtained by fitting a model on datasets  $\leq N$  and then extrapolating from it. This assesses the quality of the extrapolation. Test values are the median error rates obtained when evaluating each dataset  $>N$  on the test set of biggest available dataset



**Figure 14.** Mean-square error between error rates on the reduced test set and error rates on the full test set. This assesses the quality of the estimated errors rates, with a high quality being a necessary prerequisite for a good fit of the decay model

values appear overly optimistic compared to the fit on the largest dataset. However, this fit considers only the median error rate. When examining the best performance from each training batch of 20 networks, i.e. the network which would actually be selected for deployment, the required number of shots decreases. The “achievement value” column reports the number of

**Table 7.** Extrapolations from minimal representative dataset

Type	Problem	Shots	E	Extrapolated	Full fit	Recommendation	Achievement value
MAG	Count	1,260	2%	9,767	14,397	11,720	10,120
MAG	Discr	1,260	2%	3,572	7,705	4,286	5,059
M2	Count	948	10%	3,950	5,094	4,740	5,011
M2	Discr	948	5%	4,123	9,207	4,947	5,011

**Source(s):** The author

shots in a dataset that produces the first network to reach the desired error rate. Although the extrapolated values remain slightly optimistic, they are generally close to the actual values for practical purposes. We therefore recommend using the extrapolated prediction increased by a 20% margin. Further research with additional data is needed to provide a generalisable adjustment for this underestimation. In addition, we recommend monitoring the error rate as data acquisition progresses to detect any saturation through a gradual deviation from the power law, as observed in the case of MAG discrimination. In such cases, data acquisition can be stopped early, as additional data would provide only marginal improvement with a fixed network size.

## 10. Conclusion

In this study, we presented an approach that builds upon the original EDGAR technique to extend the instance-level classifier to ammunition type classification, while preserving its ability to learn from a weakly-labelled dataset. This solution brings significant performance improvements, simplified training, and practical deployment readiness, requiring only that the operator notes the total number of each ammunition type fired within a given time interval. By restructuring the neural network and leveraging the similarities between the counting and discrimination sub-problems, we are able to provide this additional feature for as little as a 4.2 ms inference time per shot instance, or 2.8 ms per evaluated candidate, on a 64 MHz Cortex-M4F microcontroller. On our test hardware, this results in an energy consumption of approximately 47  $\mu$ J. This implementation is ready for real-world deployment. We also provided tools and recommendations to adjust this structure for optimal discrimination performance within any fixed computational budget. We demonstrated how this alternative structure not only significantly improves performance but also simplifies the training procedure compared to a naive implementation through independent training, and introducing only easily adjustable hyperparameters for computational complexity adjustment.

We compared our approach with a widely used classical algorithm across four weapon platforms, achieving a 94% average improvement in individual error rate and perfect performance with majority voting. This enables deployment on platforms that were previously considered too complex for practical, reliable ammunition discrimination. To our knowledge, this is not only the first comprehensive description of a machine learning technique for this domain of application but also the first truly generic technique applicable to a broader range of weapon configuration discrimination tasks.

Finally, we studied how both the counting and discrimination sub-problems behave according to the size of the input dataset. From an initial sample of around 1,500 shots across sufficiently diverse weapon configurations, we are able to provide practical recommendations on the number of shots needed to reach any target error rate. We are also able to detect when the available dataset saturates the model. This improves the organizational aspects of the frequent process of dataset creation.

In the future, we aim to validate the practical performance of this technique through large-scale, real-world testing. In addition, we hope to validate the performance of this approach for the discrimination of shots fired through a suppressor. Limited testing has shown this to be feasible, with performance similar to that of the live/blank application. We also aim to further improve dataset collection by evaluating which weapon configurations are most effective for model training and, conversely, reducing or eliminating the acquisition of configurations with low added value.

With its potential for adaptability and low error rates, this approach holds promise to set a new standard for electronic shot counter capabilities, and is currently being deployed in several pilot projects. This will enable armorers worldwide to conduct tailored maintenance, as the system provides precise, automated recommendations based on ammunition counts, reducing costs, preventing incidents, and increasing operational availability.

## References

- Asbach, J. and Canty, M. (2024), "Weapon usage monitoring system having predictive maintenance and performance metrics", Patent US20240068761A1.
- Boettcher, W., Jia, X., Evans, D., Mayall, B. and Hill, S. (2024), "High precision shot detection system", Patent WO2024113022A1.
- FN Herstal (2025), "FN SmartCore® shot counter", available at: <https://fnovation.eu/products/small-arms-management/fn-smartcore/> (accessed 9 April 2025).
- Gering, A. (2020), "Device for counting live shots, blank shots and dry shots", Patent US10866048B2.
- Henighan, T., Kaplan, J., Katz, M., Chen, M., Hesse, C., Jackson, J., Jun, H., Brown, T.B., Dhariwal, P., Gray, S., Hallacy, C., Mann, B., Radford, A., Ramesh, A., Ryder, N., Ziegler, D.M., Schulman, J., Amodei, D. and McCandlish, S. (2020), "Scaling laws for autoregressive generative modeling", *arXiv preprint*, arXiv:2010.14701.
- Hestness, J., Narang, S., Ardalani, N., Diamos, G., Jun, H., Kianinejad, H., Patwary, M.M.A., Yang, Y. and Zhou, Y. (2017), "Deep learning scaling is predictable, empirically", *arXiv preprint*, arXiv:1712.00409.
- Joannes, R., Delcourt, J.-P. and Heins, P. (2010), "Device for detecting and counting shots fired by an automatic or semi-automatic firearm, and firearm equipped with such a device", Patent US7669356B2.
- Johnson, E. and Kulesza, J. (2005), "Device for collecting statistical data for maintenance of small-arms", Patent US20050155420A1.
- Kaplan, J., McCandlish, S., Henighan, T., Brown, T.B., Chess, B., Child, R., Gray, S., Radford, A., Wu, J. and Amodei, D. (2020), "Scaling laws for neural language models", *arXiv preprint*, arXiv:2001.08361.
- Khan, M.A.A.H., Welsh, D. and Roy, N. (2018), "Firearm detection using Wrist Worn Tri-Axis accelerometer signals", *2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, pp. 221-226, doi: [10.1109/PERCOMW.2018.8480345](https://doi.org/10.1109/PERCOMW.2018.8480345).
- LMT Advanced Technologies (2022), "Benefits of PAL", available at: <https://lmt-at.com/wp-content/uploads/2021/07/PAL-spec-sheet.pdf> (accessed 9 April 2025).
- Loeffler, C.E. (2014), "Detecting gunshots using wearable accelerometers", *PLoS One*, Vol. 9 No. 9, pp. 1-6, doi: [10.1371/journal.pone.0106664](https://doi.org/10.1371/journal.pone.0106664).
- Morsa, N. (2023a), "Automated system and method for a projectile launcher monitoring", Patent EP4269931A1.
- Morsa, N. (2023b), "EDGAR: embedded detection of gunshots by AI in real-time", in Guyet, T., Ifrim, G., Malinowski, S., Bagnall, A., Shafer, P. and Lemaire, V. (Eds), *Advanced Analytics and Learning on Temporal Data*, Springer International Publishing, Cham, pp. 148-166.

- Putra, I.P.E.S., Brusey, J., Gaura, E. and Vesilo, R. (2017), "An event-triggered machine learning approach for accelerometer-based fall detection", *Sensors*, Vol. 18 No. 1, p. 20, doi: [10.3390/s18010020](https://doi.org/10.3390/s18010020).
- Radetec (2024), "RISCpro installation and user manual", available at: <https://www.radetecusa.com/wp-content/uploads/2024/10/RISC-Pro-User-Manual-Version-10162024.pdf> (accessed 16 October 2024).
- Rosenfeld, J.S., Rosenfeld, A., Belinkov, Y. and Shavit, N. (2019), "A constructive prediction of the generalization error across scales", *arXiv preprint*, arXiv:1909.12673.
- Roth, M. (2024), "Heckler and Koch Bolt motion sensor system (BMSS)", *Polizeipraxis*, pp. 12-22.
- Santos, G.L., Endo, P.T., de Carvalho Monteiro, K.H., da Silva Rocha, E., Silva, I. and Lynn, T. (2019), "Accelerometer-based human fall detection using convolutional neural networks", *Sensors*, Vol. 19 No. 7, p. 1644, doi: [10.3390/s19071644](https://doi.org/10.3390/s19071644).
- Senseer (2024), "Weaponlogic by senseer", available at: [https://www.linkedin.com/posts/senseerai\\_ai-weapon-mortar-activity-7216871310302314496-Ai2m](https://www.linkedin.com/posts/senseerai_ai-weapon-mortar-activity-7216871310302314496-Ai2m) (accessed 10 December 2024).
- Staiger, M., Scheuermann, M., Kopf, J.A., Gebert, D. and Rimpf, D. (2023), "Firearm analysis devices", Patent US11802747B2.
- Ufer, R. and Brinkley, K.L. (2014), "Self calibrating weapon shot counter", Patent US8826575B2.
- Weiss, I., Hyden, I. and Ami, M. (2021), "Device system and method for projectile launcher operation monitoring", Patent US20210199400A1.
- Welsh, D., Faridee, A.Z.M. and Roy, N. (2021), "Hybrid distance-based framework for classification of embedded firearm recoil data", 2021 IEEE International Conference on Pervasive Computing and Communications Workshops and Other Affiliated Events (PerCom Workshops), pp. 50-55. doi: [10.1109/PerComWorkshops51409.2021.9431020](https://doi.org/10.1109/PerComWorkshops51409.2021.9431020).
- Welsh, D. and Roy, N. (2017), "Smartphone-based mobile gunshot detection", 2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), pp. 244-249, doi: [10.1109/PERCOMW.2017.7917566](https://doi.org/10.1109/PERCOMW.2017.7917566).

#### Corresponding author

Nathan Morsa can be contacted at: [nathan.morsa@uliege.be](mailto:nathan.morsa@uliege.be)