

G-Induced Loss of Consciousness Prediction Using a Support Vector Machine

Nobuhiro Ohrui; Yuji Iino; Koichiro Kuramoto; Azusa Kikukawa; Koji Okano; Kunio Takada; Tetsuya Tsujimoto

- INTRODUCTION:** Gravity-induced loss of consciousness (G-LOC) is a major threat to fighter pilots and may result in fatal accidents. The brain has a period of 5–6 s from the onset of high $+G_z$ exposure, called the functional buffer period, during which transient ischemia is tolerated without loss of consciousness. We tried to establish a method for predicting G-LOC within the functional buffer period by using machine learning. We used a support vector machine (SVM), which is a popular classification algorithm in machine learning.
- METHODS:** The subjects were 124 flight course students. We used a linear soft-margin SVM, a nonlinear SVM Gaussian kernel function (GSVM), and a polynomial kernel function, for each of which 10 classifiers were built every 0.5 s from the onset of high $+G_z$ exposure (Classifiers 0.5–5.0) to predict G-LOC. Explanatory variables used for each SVM were age, height, weight, with/without anti-G suit, $+G_z$ level, cerebral oxyhemoglobin concentration, and deoxyhemoglobin concentration.
- RESULTS:** The performance of GSVM was better than that of other SVMs. The accuracy of each classifier of GSVM was as follows: Classifier 0.5, 58.1%; 1.0, 54.8%; 1.5, 57.3%; 2.0, 58.1%; 2.5, 64.5%; 3.0, 63.7%; 3.5, 65.3%; 4.0, 64.5%; 4.5, 64.5%; and 5.0, 64.5%.
- CONCLUSION:** We could predict G-LOC with an accuracy rate of approximately 65% from 2.5 s after the onset of high $+G_z$ exposure by using GSVM. Analysis of a larger number of cases and factors to enhance accuracy may be needed to apply those classifiers in centrifuge training and actual flight.
- KEYWORDS:** gravity-induced loss of consciousness, machine learning, support vector machine.

Ohruai N, Iino Y, Kuramoto K, Kikukawa A, Okano K, Takada K, Tsujimoto T. *G-induced loss of consciousness prediction using a support vector machine. Aerosp Med Hum Perform. 2024; 95(1):29–36.*

Gravity-induced loss of consciousness (G-LOC) is a major threat to high performance fighter pilots. A U.S. Air Force study showed that G-LOC-related accidents occurred 25.1 times per million flights during the period from 1982 to 2002 and that 3.6% of those accidents were fatal.⁷ The rate of occurrence of G-LOC in aircrew in other countries has been reported to be 8.2–20.1%.^{3,4,14} G-LOC is caused by exposure to high and sustained levels of head-to-foot G force ($+G_z$) that reduce cerebral blood flow, resulting in a reduction in oxygen supply to the brain. To prevent G-LOC, centrifuge training for high performance fighter pilots is conducted in many countries. The Japan Air Self-Defense Force (JASDF) requires its pilot candidates to undergo centrifuge training as a part of the initial training, and its fighter pilot trainees are required to undergo centrifuge training as a part of the advanced training.¹⁰ During centrifuge training, trainers monitor the pilot's facial

expressions, loss of response, and loss of postural tone to detect G-LOC. However, during actual flight, it is impossible to detect G-LOC because there is no monitoring of the pilot's facial expressions, loss of response, or loss of postural tone. Therefore, G-LOC is only detected during centrifuge training. Although there have been many studies on the detection of onset of G-LOC using physiological monitoring without monitoring of

From the Aeromedical Laboratory and the Aero Safety Service Group, Japan Air Self-Defense Force, Sayama, Saitama, Japan.

This manuscript was received for review in May 2023. It was accepted for publication in October 2023.

Address correspondence to: Nobuhiro Ohruai, Ph.D., 2-3 Inariyama, Sayama, Saitama 350-1324, Japan; oruulucky@gmail.com.

Reprint and copyright © by the Aerospace Medical Association, Alexandria, VA.

DOI: <https://doi.org/10.3357/AMHP.6301.2024>

facial expressions, loss of response, and loss of postural tone, an onset of G-LOC detection system has not yet been developed.

Near-infrared spectroscopy (NIRS) is an easy-to-use and noninvasive method for continuous real-time monitoring of cerebral cortical oxygenation. NIRS has been used by the JASDF to try to predict G-LOC during centrifuge training.^{5,11} It is thought that G-LOC occurs when the cerebral oxyhemoglobin concentration (OxyHb) measured by NIRS falls below a certain level that maintains brain functions. However, the extent of decrease in OxyHb that leads to G-LOC is not known and there are, therefore, limitations of NIRS for G-LOC detection and prediction. We have established a formula for predicting G-LOC by logistic regression analysis using rate of change from maximum to minimum values of OxyHb and body mass index as explanatory variables.¹¹ The formula for predicting G-LOC showed sensitivity, specificity, and accuracy of 67.6%, 81.4%, and 79.5%, respectively. However, this formula has a few problems: 1) it cannot predict G-LOC until just before the occurrence of G-LOC; and 2) the rate of change from maximum to minimum values of OxyHb must be calculated in advance and G-LOC, therefore, cannot be predicted in real time. For these reasons, we are trying to establish a more effective way to predict G-LOC in real time.

The brain has the ability to tolerate transient ischemia without loss of consciousness during high $+G_z$ exposure. The period of tolerance has been termed the “functional buffer period (FBP)” and lasts for approximately 5–6 s against any $+G_z$ exposure.^{8,16} Therefore, it has been thought that G-LOC does not occur during the FBP. Prediction of G-LOC as soon as possible from the onset of high $+G_z$ exposure during the FBP would be extremely useful as a G-LOC countermeasure. We previously analyzed OxyHb during the FBP (5 s after the onset of high $+G_z$ exposure) to predict whether G-LOC would occur.¹² Unfortunately, despite analyzing the results using a variety of statistical analyses, we were unable to find a specific pattern that predicted G-LOC.

There have been many developments in machine learning techniques that use computers to extract specific patterns from various events. Statistical analysis focuses on explaining or interpreting the nature of data, whereas machine learning is a method that focuses on more accurate prediction and classification. Machine learning enables the analysis of large volumes of complex data that would be impossible for humans and enables pattern recognition and generation of highly accurate predictive models that cannot be found using statistical analysis. There are three main machine learning algorithms: supervised, unsupervised, and reinforcement learning. Supervised learning uses input-output pairs or labeled data to train a model to produce a function. It enables the model to predict future outcomes from being trained on past data. A support vector machine (SVM) is one of the most widely used supervised learning models that can be applied to problems such as classification and regression, and it is a learning method that classifies a population into two classes and determines which one the unknown data belong to. In this study, we attempted to predict G-LOC within 5 s after the onset of

high $+G_z$ by using an SVM with age, height, weight, with/without anti-G suit, $+G_z$ level, OxyHb, and cerebral deoxyhemoglobin concentration (DeoxyHb) as explanatory variables.

METHODS

Subjects

Approval for this study was obtained from the JASDF Human Ethics Committee, the Aeromedical Laboratory. The subjects were advanced flight course trained students who participated in JASDF centrifuge training from 2008 to 2012. An informed consent form was signed by each of the subjects. We classified subjects into G-LOC occurrence or nonoccurrence according to their centrifuge training results. There was an imbalance between the numbers of subjects with G-LOC occurrence and without G-LOC occurrence. Therefore, subjects with G-LOC occurrence and without G-LOC occurrence were sampled an equal number using a random sampling. The subjects were 124 healthy men [mean (SD) age, 24.3 (1.8) yr]. The numbers of the subjects with and those without G-LOC occurrence during the $+G_z$ exposure conditions were 17 in each group for $+6 G_z$, 31 in each group for $+7 G_z$, and 14 in each group for $+8 G_z$.

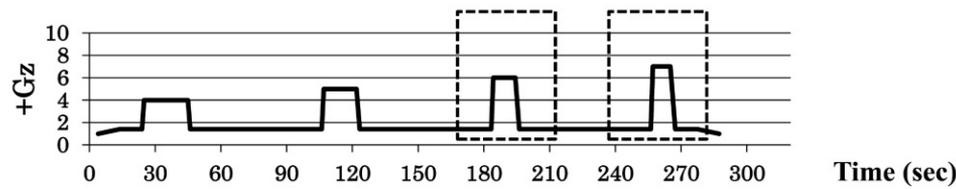
Procedure

The current study involved prediction of G-LOC during centrifuge training. The JASDF Aeromedical Laboratory Centrifuge (radius of 7.6 m) at Iruma Air Base, Saitama, Japan, was used for the centrifuge training. The JASDF has eight centrifuge profiles for advanced flight course training (Profiles 1–8).¹⁰ G-LOC occurs more frequently in Profiles 2 and 6 (**Fig. 1**) among the centrifuge profiles. In Profile 2, after bringing the centrifuge to a $+1.4-G_z$ idle run condition, subjects without an anti-G suit were first exposed to $+4 G_z$ at an onset rate of $+1 G_z \cdot s^{-1}$ for 15 s and then to $+5 G_z$ for 10 s, $+6 G_z$ for 8 s, and $+7 G_z$ for 8 s. There was a 1-min rest plateau of the $+1.4-G_z$ idle run condition between the G pulses. In Profile 6, after bringing the centrifuge to a $+1.4-G_z$ idle run condition, subjects with an anti-G suit were exposed to sustained $+8 G_z$ with onset rates of $+6 G_z \cdot s^{-1}$ for 15 s. All subjects performed an anti-G straining maneuver during $+G_z$ exposures in Profiles 2 and 6. In this study, the following centrifuge conditions were analyzed: 1) during $+6 G_z$ in Profile 2 ($+6 G_z$, onset rate of $+1 G_z \cdot s^{-1}$, without an anti-G suit) ($+6 G_z$); 2) during $+7 G_z$ in Profile 2 ($+7 G_z$, onset rate of $+1 G_z \cdot s^{-1}$, without an anti-G suit) ($+7 G_z$); and c) during $+8 G_z$ in Profile 6 ($+8 G_z$, onset rate of $+6 G_z \cdot s^{-1}$, with an anti-G suit) ($+8 G_z$).

During each of the $+G_z$ exposure conditions, OxyHb and DeoxyHb were recorded by using a NIRO-150G near-infrared spectrophotometer (Hamamatsu Photonics K.K., Shizuoka, Japan). The methods used have been described previously.⁵ The optodes (light source and light detector) in a specialized rubber holder were set at a constant distance of 4.0 cm apart on the left forehead with avoidance of temporal muscle regions. Data were logged every 0.5 s during the G exposure. Baseline values of OxyHb and DeoxyHb were obtained as average values over a

Profile 2

- 13 degree seatback angle
- without anti-G suits
- Onset rate: 1G/sec
- Plateau +G_z level (sustained time): 4 (15 sec), 5(10 sec), 6 (8 sec) and 7 (8 sec)

**Profile 6**

- 13 degree seatback angle
- with anti-G suits
- Onset rate: 6G/sec
- Plateau +G_z level (sustained time): 8 (15 sec)

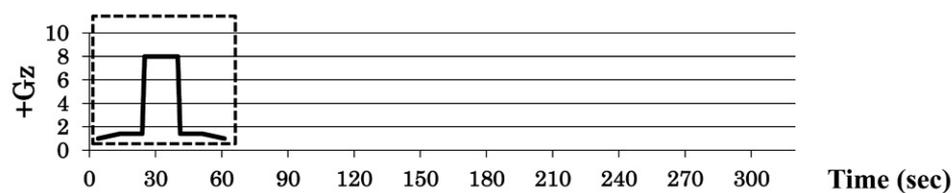


Fig. 1. Profiles of centrifuge training in the Aeromedical Laboratory, Japan Air Self-Defense Force (advanced training 2 and 6). In this study, the following centrifuge training conditions were analyzed: A) during +6 G_z in Profile 2 (+6 G_z, onset rate of 1 G_z · s⁻¹, without an anti-G suit); B) during +7 G_z in Profile 2 (+7 G_z, onset rate of +1 G_z · s⁻¹, without an anti-G suit); and C) during +8 G_z in Profile 6 (+8 G_z, onset rate of +6 G_z · s⁻¹, with an anti-G suit) (circled in dashed lines).

30-s period before +G_z exposure in each profile. During this period, the subjects were still and quiet on the seat in the centrifuge gondola waiting to run.

An SVM is a widely used supervised machine learning algorithm. It can generalize between two different group classifications. SVMs include linear SVMs and nonlinear SVMs.

A linear SVM is used for linearly separable data. For example, as shown in **Fig. 2**, suppose that the input space contains two classes of training data, black circles and white circles. If the training data are linearly separable, then the decision boundary separating hyperplanes of the two classes will be represented by the following equation. Superscript T denotes the transpose of the vector.

$$w^T x + b = 0.$$

The combination of w and b that defines the decision boundary is not uniquely determined but exists infinitely. In other words, SVM training can be rephrased as the problem of finding w and b that will give the best generalization performance from the training data. However, a classifier must not only be able to correctly identify known training data, but also have the generalization capability to correctly identify against unknown data. Therefore, the distance between the decision boundary and the nearest known training data point from the decision boundary is considered to account for the generalization performance for unknown data. This distance is called the margin. Choosing a decision boundary with a large margin is more likely to correctly

classify unknown data; that is, SVM is a method for defining the decision boundary from given training data in a way that maximizes the margin. The solution of the problem for w and b , such that the margin is maximized, can be attributed to the optimization problems of a convex quadratic function, subject to linear constraints, and the uniqueness of the solutions holds.

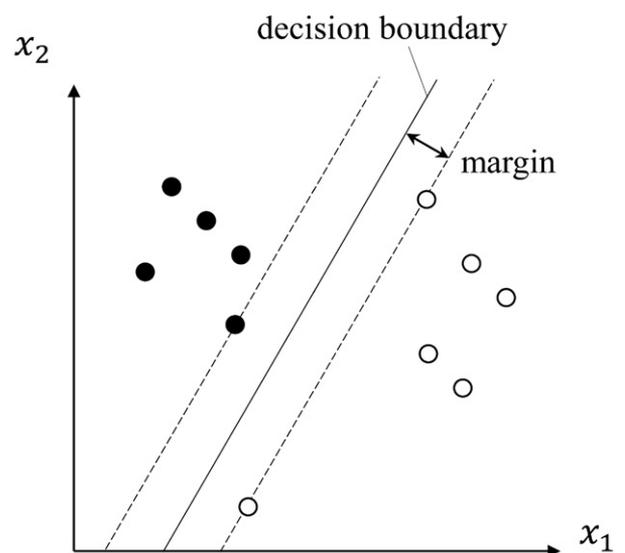


Fig. 2. Test data in the feature space and the decision boundary. The margin is the distance between the decision boundary and the closest data points.

The vector from the decision boundary to the nearest training data point thus obtained can be considered as supporting the optimal decision boundary. These are called support vectors and the algorithm is, therefore, called SVM. The solid line in Fig. 2 shows the decision boundary for generalization between two different group classifications of the training data. The dashed lines on both sides indicate the range from the decision boundary. SVM classification includes hard margin SVM, which finds the decision boundary under conditions that allow zero errors for the training data, and soft margin SVM, which finds the decision boundary under constraints that tolerate some errors using an added penalty term (box constraints). In practice, soft margin SVMs are often used because the training data may not be completely linearly separated and the generalization performance of hard margin SVMs might be reduced due to overfitting.

In Fig. 2, the training data could be linearly separated and a decision boundary separating the two classes could therefore be obtained. The soft-margin SVM can also be used to find a linear decision boundary even if it is affected by noise or other factors. However, the linear SVM using a linear decision boundary cannot provide sufficient generalization

performance when the data of both classes are complexly distributed in the input space. In the nonlinear SVM, a nonlinear function is used for mapping the data of the input space to a higher dimensional feature space in which a linear separating hyperplane can be found. The decision boundary is then defined in this feature space. In Fig. 3A, it is not possible to define the decision boundary separating the hyperplanes of the two classes. As a solution to this problem, the feature data sets are nonlinearly transformed into a higher dimension. In this space, the separating hyperplane may be definable (Fig. 3B). The separating hyperplane found in the new space corresponds to a nonlinear decision boundary in the original space (Fig. 3C). Nonlinear functions that are often used to map training data are Gaussian and polynomial kernel functions. Nonlinear SVMs have high generalization performance for data that cannot be linearly separated.

In the present study, three SVM algorithms were used: linear soft-margin SVM (LSVM), nonlinear SVM Gaussian kernel function (GSVM), and polynomial kernel function (PSVM), for each of which 10 classifiers were built every 0.5 s from the onset of high $+G_z$ exposure (Classifiers 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, and 5.0).

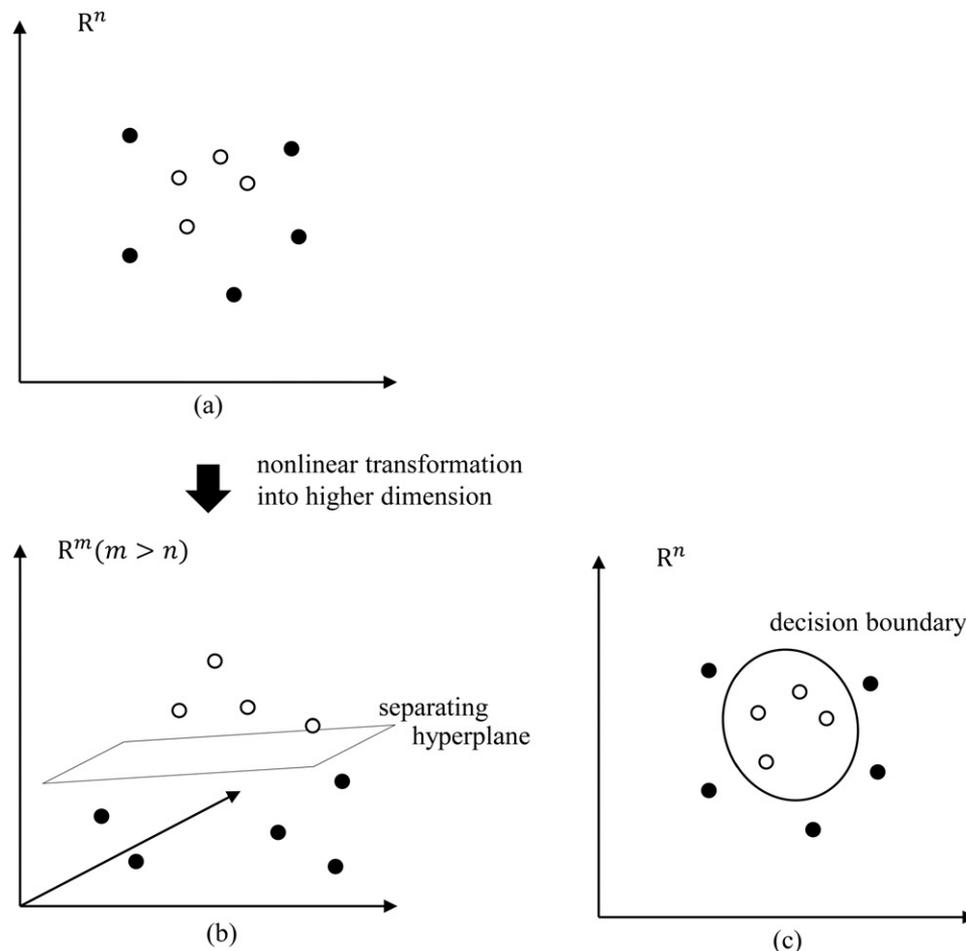


Fig. 3. Principle of a nonlinear SVM. A) Test data sets in the input space are not linearly separable. B) Feature data sets are nonlinearly transformed into a higher dimension. In this space, the separating hyperplane may be definable. C) Decision boundary in the input space.

SVM classifiers were built with G-LOC occurrence or non-occurrence as response variables. Explanatory variables for the SVM were age (yr), height (cm), weight (kg), with or without an anti-G suit (categorical data), $+G_z$ level at 0–5.0s, OxyHb at 0.0–5.0s, and DeoxyHb at 0.0–5.0s. The number of items in each classifier was as follows: Classifier 0.5s, 10; Classifier 1.0, 13; Classifier 1.5, 16; Classifier 2.0, 19; Classifier, 2.5, 22; Classifier 3.0, 25; Classifier 3.5, 28; Classifier 4.0, 31; Classifier 4.5, 34; and Classifier 5.0, 37. For example, in Classifier 1.0, the 13 items were: 1) age; 2) height; 3) weight; 4) with or without an anti-G suit; 5) G values at 0, 6) at 0.5, and 7) at 1.0s; OxyHb at 8) 0, 9) 0.5, and 10) 1.0s; and DeoxyHb at 11) 0, 12) 0.5, and 13) 1.0s. There was a $+1.4-G_z$ idle run condition in exposure profiles. Thus, the $+G_z$ level was taken as a value every 0.5 s, starting at $+1.4 G_z$ ($0s = +1.4 G_z$).

We consider it is technically feasible to determine the rates of changes in OxyHb and DeoxyHb in real time. However, we are considering a method using commercial products. The NIRS system currently in our possession does not have the capability to calculate rates of change in real time. In addition, to the best of our knowledge, there is currently no commercially available NIRS system that is capable of calculating rates of change in real time. In this study, we did not use calculated data such as the rates of changes in OxyHb and DeoxyHb and we used data as output from NIRS, such as OxyHb and DeoxyHb, because we want to make G-LOC forecasting in real time in the future.

In the present study, a leave-one-out cross validation was used to assess the performance of the classifiers. In the leave-one-out cross validation, one example is extracted from the sample group and used as test data, while the rest is used as training data. This is a method in which all data are repeatedly

validated so that they become test data one by one. This study was conducted on 124 data points (123 data points were used as training data and the remaining 1 data point was used as test data) for a total of 124 times (Fig. 4).

The performance of the classifiers built by the training data was evaluated using the test data. We evaluated the performance of the classifiers from each of the perspectives of precision, recall, and accuracy. Table I shows a confusion matrix. When the classifiers built in this study were given unknown test data, the prediction output by the classifiers was one of the following four types: 1) true positive in cases in which the classifiers predicted G-LOC would occur and G-LOC actually did occur; 2) false positive in cases in which the classifiers predicted G-LOC would occur, but G-LOC did not actually occur; 3) true negative in cases in which the classifiers predicted G-LOC would not occur and G-LOC actually did not occur; and 4) false negative in cases in which the classifiers predicted G-LOC would not occur, but G-LOC actually did occur. Precision was defined as the percentage of subjects who experienced G-LOC in subjects who were predicted to experience G-LOC. Recall was defined as the percentage of subjects who were predicted to experience G-LOC in subjects who experienced G-LOC. Accuracy was defined as the percentage of all subjects who were correctly predicted to experience or not to experience G-LOC.

The performance of classifiers depends on the hyperparameters. For this reason, optimization of the hyperparameters was repeated 500 times for each classifier using Bayesian optimization to maximize using accuracy as an index, and the optimized results were used as the respective performance evaluation. The respective optimized hyperparameters are as follows: LSVM, cost of regularization; GSVM, cost of regularization and kernel coefficient parameter gamma; and PSVM, cost of regularization, kernel coefficient parameter gamma, and degree of polynomial.² The training data were scaled beforehand by normalization and standardization. This ensures that they would not be affected by differences in the scale of the explanatory variables such as the magnitude of the value.

The permutation feature importance (PFI) is a method for evaluating how much an explanatory variable contributes to the predictive accuracy of a model. PFI is determined by the importance of each explanatory variable depending on how much the prediction accuracy drops; in other words, increase in the prediction error (PE) of the model, when a single explanatory variable is randomly shuffled, which breaks the relationship between the explanatory variable and the true outcome. We used PFI to evaluate the explanatory variables in high accuracy classifiers with the highest performance in the

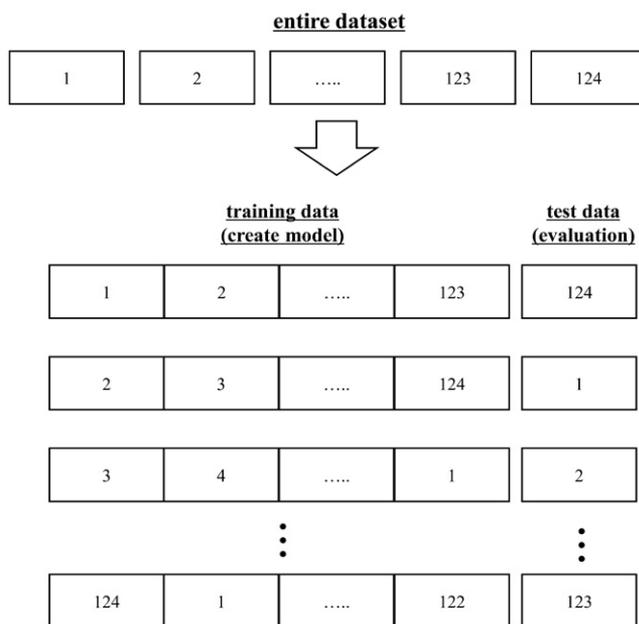


Fig. 4. Leave-one-out cross validation. This study was conducted on 124 data points: 123 data points were used as training data and the remaining 1 data point was used as test data.

Table I. Confusion Matrix.

PREDICTED G-LOC	OBSERVED G-LOC	
	PRESENT	ABSENT
Present	True Positive (TP)	False Positive (FP)
Absent	False Negative (FN)	True Negative (TN)

Precision = $TP/(TP+FP)$; Recall = $TP/(TP+FN)$; Accuracy = $(TP+TN)/(TP+FP+TN+FN)$.

three SVM algorithms used in this study. The classifiers were built and their performance evaluation and optimization were performed using Python 3.9¹⁵ and its machine learning package scikit-learn 1.1.1.¹³

Statistical Analysis

The results are expressed as means \pm SD. Data on age, height, and weight in the G-LOC group and Non-G-LOC group in +6 G_z , +7 G_z , and +8 G_z were compared using Student's *t*-test. We also used a two-way mixed design analysis of variance to examine changes in OxyHb and DeoxyHb in +6 G_z , +7 G_z , and +8 G_z , with or without G-LOC (between subjects) and G-loading time (within subjects) as two factors. The equivariance of within-subject factors was tested for Mauchly's sphericity, and if sphericity could not be assumed, the degrees of freedom were adjusted for Greenhouse-Geisser. Multiple comparison tests were performed using the Bonferroni method. The levels of each factor are: 1) G-LOC occurrence: 2 levels (G-LOC group and Non-G-LOC group); and 2) + G_z loading time: 11 levels [0–5 s (0.5-s intervals)]. Differences were considered significant at $P < 0.05$ in all analyses. All statistical analyses were performed using SPSS (version 24.0, Chicago, IL, USA).

RESULTS

The times from onset of high + G_z load to onset of G-LOC were 10.4 ± 1.4 s (minimum–maximum: 8.5–12.5) at +6 G_z , 10.6 ± 1.8 s (6.0–13.5) at +7 G_z , and 7.8 ± 1.1 s (6.0–9.5) at +8 G_z . There were no significant differences (n.s.) in age (+6 G_z : G-LOC group, 24.4 ± 0.5 vs. Non-G-LOC, 24.8 ± 0.5 ; $t = -0.624$, $P = 0.537$; +7 G_z : 24.1 ± 0.3 vs. 24.0 ± 0.3 ; $t = 0.306$, $P = 0.761$; +8 G_z : 23.8 ± 0.4 vs. 25.1 ± 0.5 ; $t = -1.948$, $P = 0.063$), height (+6 G_z : 173.1 ± 1.0 vs. 173.5 ± 1.1 ; $t = -0.269$, $P = 0.789$; +7 G_z : 172.0 ± 0.9 vs. 170.6 ± 0.9 ; $t = 1.149$, $P = 0.255$, +8 G_z : 173.1 ± 1.6 vs. 173.4 ± 1.9 ; $t = -0.114$, $P = 0.910$), or weight (+6 G_z : 65.0 ± 1.2 vs. 67.8 ± 1.4 ; $t = -1.480$, $P = 0.149$; +7 G_z : 65.1 ± 1.3 vs. 64.3 ± 1.1 ; $t = 0.483$, $P = 0.631$; +8 G_z : 66.0 ± 2.7 vs. 69.5 ± 2.4 ; $t = -0.972$, $P = 0.340$) between the G-LOC group and Non-G-LOC group in +6 G_z , +7 G_z , and +8 G_z , respectively.

For OxyHb, there was no interaction between G-LOC occurrence and G-loading time in +6 G_z [$F(2.036, 65.164) = 0.577$, $P = 0.567$] or +7 G_z [$F(2.555, 153.279) = 0.161$, $P = 0.897$]. The main effect of G-LOC occurrence was significant in +6 G_z [$F(1, 32) = 4.175$, $P = 0.049$] and the G-LOC group had lower OxyHb than that in the Non-G-LOC group, but the difference was not significant in +7 G_z [$F(1, 60) = 0.396$, $P = 0.532$]. The main effect of G-loading time was significantly different in OxyHb in +6 G_z [$F(2.036, 65.164) = 53.912$, $P < 0.001$] and +7 G_z [$F(2.555, 153.279) = 112.730$, $P < 0.001$]. In the results of multiple comparison tests, OxyHb was significantly lower at all time points from 1.5 s onward compared to that at 0 s ($P < 0.05$) in +6 G_z and was significantly lower at all time points from 2.0 s compared to that at 0 s ($P < 0.05$) in +7 G_z . There was a significant interaction between G-LOC occurrence and G-loading time in OxyHb in +8 G_z [$F(1.717, 44.637) = 5.465$, $P = 0.010$]. Because a significant

interaction was observed, a comparison of OxyHb between the G-LOC group and the Non-G-LOC group at each G-loading time showed that OxyHb was significantly higher in the G-LOC group than in the Non-G-LOC group at 0 s [$F(1, 26) = 4.270$, $P = 0.049$], with no significant differences at other G-loading times. The main effect of G-loading time was significantly different in OxyHb in the G-LOC group and the Non-G-LOC group at +8 G_z . The G-LOC group at +8 G_z had significantly lower OxyHb at all other G-loading times compared to that at 0 s ($P < 0.05$). In the Non-G-LOC in +8 G_z group, OxyHb was significantly lower at all + G_z -loading times after 1.0 s than at 0 s ($P < 0.05$).

For DeoxyHb, there was no interaction between G-LOC occurrence and G-loading time at +6 G_z [$F(1.774, 56.776) = 0.273$, $P = 0.736$], +7 G_z [$F(2.328, 139.664) = 0.345$, $P = 0.741$], and +8 G_z [$F(2.307, 59.987) = 1.796$, $P = 0.170$]. The main effect of G-LOC occurrence was significant in +8 G_z [$F(1, 26) = 17.382$, $P < 0.001$] and the G-LOC group had higher DeoxyHb than that in the Non-G-LOC group, but the difference was not significant at +6 G_z [$F(1, 32) = 0.285$, $P = 0.597$] or +7 G_z [$F(1, 60) = 0.386$, $P = 0.537$]. The main effect of G-loading time was significantly different in DeoxyHb at +6 G_z [$F(1.774, 56.776) = 16.515$, $P < 0.001$], +7 G_z [$F(2.328, 139.664) = 22.166$, $P < 0.001$], and +8 G_z [$F(2.307, 59.987) = 20.108$, $P < 0.001$]. In the results of multiple comparison tests, DeoxyHb was significantly higher at all time points from 2.0 s onward compared to that at 0 s at +6 G_z and +7 G_z ($P < 0.05$) and was significantly higher at all time points compared to that at 0 s in +8 G_z ($P < 0.05$).

Table II shows precision, recall, and accuracy for Classifiers 0.5–5.0 using LSVM, GSVM, and PSVM. For LSVM, the accuracy exceeded 60% from Classifier 3.0 onwards. It leveled off between Classifiers 3.0 and 5.0 and ranged from 60.5–64.5%. The precision of Classifiers 3.0–5.0 ranged from 61.0–64.4% and recall ranged from 58.1–66.1%. GSVM showed superior and more consistent generalization results among the three types of SVM tested in this study. The accuracy exceeded 60% from Classifier 2.5 onwards. It leveled off between Classifiers 2.5 and 5.0 and ranged from 63.7–65.3%. The precision of Classifiers 2.5–5.0 ranged from 62.7–65.1% and recall ranged from 62.9–71.0%. PSVM showed accuracy of over 60% at Classifier 2.0, the earliest stage of the three types of SVM built into this study. However, it was not stable for later classifiers, with accuracy from 57.3–63.7% for Classifiers 2.0–5.0. The precision of Classifiers 2.0–5.0 ranged from 56.0–62.7% and recall ranged from 59.7–69.4%.

The importance of explanatory variables in Classifiers 2.5–5.0 with an accuracy of over 60% among GSVMs that showed superior generalization results in this study was evaluated using PFI. The results showed that age (PE; 0.17 ± 0.01), height (PE; 0.24 ± 0.00) and immediate OxyHb (e.g., if Classifier 2.5, OxyHb at 2.5 s) (PE; 0.25 ± 0.03) were more important than the other explanatory variables (PE; 0.01 ± 0.01).

DISCUSSION

G-LOC countermeasures such as the anti-G suit and anti-G straining maneuvers have reduced the number of G-LOC cases,

Table II. Classifier Performance.

	CLASSIFIERS									
	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Linear Soft-Margin SVM										
Precision (%)	52.2	54.0	55.6	54.0	54.2	61.0*	63.3*	62.3*	64.4*	64.1*
Recall (%)	56.5	54.8	56.5	54.8	51.6	58.1	61.3*	61.3*	61.3*	66.1*
Accuracy (%)	52.4	54.0	55.6	54.0	54.0	60.5*	62.9*	62.1*	63.7*	64.5*
Nonlinear SVM Gaussian Kernel Function										
Precision (%)	61.4*	54.7	57.4	58.3	62.9*	62.7*	65.1*	63.2*	64.5*	65.0*
Recall (%)	43.5	56.5	56.5	56.5	71.0*	67.7*	66.1*	69.4*	64.5*	62.9*
Accuracy (%)	58.1	54.8	57.3	58.1	64.5*	63.7*	65.3*	64.5*	64.5*	64.5*
Nonlinear SVM polynomial kernel function										
Precision (%)	44.6	55.0	54.2	62.7*	61.5*	62.5*	62.3*	59.2	59.7	56.0
Recall (%)	46.8	53.2	83.9*	59.7	64.5*	64.5*	69.4*	67.7*	64.5*	67.7*
Accuracy (%)	44.4	54.8	56.5	62.1*	62.1*	62.9*	63.7*	60.5*	60.5*	57.3

*Indicates > 60%.

SVM = support vector machine.

but G-LOC still occurs in some cases. Many efforts have been made to try to detect and predict G-LOC in flight. However, technology is not yet established. It has been suggested that G-LOC does not occur during the FBP that lasts for approximately 5–6 s during high + G_z exposure.^{8,16} In this study, G-LOC occurred on average 7.8–10.6 s from the onset of high + G_z exposure and occurred at 6 s or more in all cases. The time from onset of high + G_z load to onset of G-LOC at +8 G_z was 7.8 s and is more reflective of what pilots will experience in the real world as they will experience rapid G onset. It would be extremely useful for aviation safety and accident prevention if G-LOC could be predicted within the several seconds during which brain function is maintained. We aim to apply the classifiers to trigger the activation of a G-LOC warning light and/or sound and automatic recovery systems during centrifuge training and in flight in the future. The Automatic Ground Collision Avoidance System, developed by Lockheed Martin and others, provides a warning on the heads-up display if the system predicts a collision with the ground and automatically activates the system if there is no pilot response.⁶ For example, integration into such a device could alert the pilot before the pilot becomes unresponsive, which might allow the pilot to perform avoidance maneuvers (e.g., reduce high + G_z exposure) or allow automatic avoidance maneuvers more quickly, albeit for a short period of time (e.g., if the prediction of G-LOC occurs at 2.5 s and G-LOC occurs 5 s later at the 7.8-s mark for rapid G onset).

In the present study, three types of SVMs were used: LSVM, GSVM, and PSVM, for each of which 10 classifiers every 0.5 s from the onset of high + G_z exposure were built. GSVM performed the best in terms of accuracy. The accuracies were as follows: Classifier 2.5, 64.5%; Classifier 3.0, 63.7%; Classifier 3.5, 65.3%; Classifier 4.0, 64.5%; Classifier 4.5, 64.5%; and Classifier 5.0, 64.5%. These results suggest that it may be possible to predict whether G-LOC will occur with an accuracy rate of approximately 65% at 2.5 s from the onset of high + G_z exposure. On the other hand, as the number of items of explanatory variables increased over time, accuracy is expected to improve. However, it leveled off up to Classifier 5.0 s. A very common way to evaluate the performance of classifiers in an SVM is to measure precision, recall, and accuracy. The precision is the

percentage of G-LOC predictions that came true, recall is the percentage of G-LOCs predicted without oversight, and accuracy is the percentage of correct predictions of whether G-LOC occurred. If we were using statistical analysis to predict this, we would have *P*-values, correlation coefficients, etc. to help us determine if in fact our predictions were accurate and significant. In contrast, determination of good or bad precision, recall, and accuracy in SVMs is subjective. We believe, however, that 50% may be one baseline in this study. This is because when the data are perfectly balanced, as in this study in which subjects for the G-LOC and Non-G-LOC groups were sampled in equal numbers, always predicting “G-LOC” or “Non-G-LOC” will yield an accuracy of 50%. In this case, if the classifiers learn nothing useful, the accuracy is likely to be around 50% and anything above 50% is considered better than a random guess. The results in this study exceed this 50%. However, accuracy, repeatability, and precision of 65% are not good enough for operational purposes and there is a need for further improvement. There is a tradeoff between precision and recall. If one increases, the other decreases. Precision and recall cannot be increased simultaneously. If there is concern about false positives, precision should be emphasized, and if there is concern about false negatives, recall should be emphasized. When considering applications such as triggering automatic recovery systems, accuracy of around 65% for the classifiers would result in many malfunctions. This would make it impractical. It is considered important from a safety perspective to prevent oversight of G-LOC in centrifuge training and flight training. In terms of recall, as with accuracy, G-LOCs were predicted by Classifier 2.5 without oversight at approximately 65%, and this remained unchanged up to Classifier 5.0. In principle, the goal is to increase the accuracy and lead to practical applications, but if it is permissible for the device to malfunction to some extent during centrifuge training and flight training, such as a warning tone, a classifier that emphasizes recall may be an option. In any case, the performance of the classifier needs to be improved further for practical applications. The performance of an SVM depends on the amount of training data. Further accumulation of training data may improve performance. There is also the expectation that performance will be improved in the future by

further increases in the number of explanatory variables for the SVM. G-LOC results from a reduced blood supply from the heart to the brain. Therefore, the performance of the classifier could be enhanced by adding physiological indicators such as an electroencephalogram, blood pressure, and heart rate to the explanatory variables. In addition, there is scope to consider whether machine learning other than SVMs, such as Random Forests, can be applied.

There has been much discussion about the associations of physiological variables such as age, height, weight, and risk with G-LOC.⁹ In the present study, statistical analysis did not show an association between age, height, or weight and the risk of G-LOC, but the SVM suggested that age and height were relatively important explanatory variables compared to the other explanatory variables. Also, in OxyHb and DeoxyHb, statistical analysis did not show a constant trend in +6 G_z, +7 G_z, and +8 G_z, but the SVM suggested that immediate OxyHb was a relatively important explanatory variable compared to the other explanatory variables. These results might indicate that an SVM has the potential to find relationships between physiological variables and G-LOC that are not found by statistical analysis and thus predict G-LOC.

This study was conducted by postprocessing and analyzing previously acquired data. In the future, easy attachment of measurement equipment and real-time data analysis will be required for centrifuge training and actual flight applications. Achievement of this goal will require reduction in the size and weight and ruggedization of measuring equipment as well as improved performance such as greater computer processing speed. Various wearable devices such as heart rate monitors and electrocardiographs in wristband and clothing forms have been developed in recent years. Moreover, the U.S. Air Force has developed an integrated cockpit sensing system that enables real-time monitoring of the pilot's physiological indicators, such as blood oxygen saturation, blood perfusion, heart rate, heart rate variability, estimated core temperature, skin temperature, respiration rate, and work of breathing, and the cockpit environment during actual flight.¹ We look forward to the development and popularization of equipment that is wearable and stable in the harsh aviation environment and which enables analysis in real time.

In conclusion, three types of SVM algorithms were used: LSVM, GSVM, and PSVM, for each of which 10 classifiers were built every 0.5 s from the onset of high +G_z exposure (Classifiers 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, and 5.0) and their performance was evaluated. GSVM showed the most stable performance of the three algorithms tested in this study, with accuracy rates of 63.7–65.3% after Classifier 2.5. Analysis of a larger number of cases and factors to enhance accuracy may be needed to apply those classifiers in centrifuge training and actual flight.

ACKNOWLEDGMENTS

We are greatly indebted to the staff of the Centrifuge Training Section of the JASDF Aeromedical Laboratory and Shunsuke Hayashi of the Human Engineering Section of the JASDF Aeromedical Laboratory for their technical

support with data analysis methods. We would like to thank S.E.S. Translation and Proofreading Services for English language editing.

Financial Disclosure Statement: The authors have no competing interests to declare.

Authors and Affiliations: Nobuhiro Ohruai, Ph.D., Koichiro Kuramoto, Ph.D., Azusa Kikukawa, Ph.D., Koji Okano, M.A., Kunio Takada, M.D., Ph.D., and Tetsuya Tsujimoto, M.D., Aeromedical Laboratory, and Yuji Iino, M.Eng., Aero Safety Service Group, Japan Air Self-Defense Force, Saitama, Japan.

REFERENCES

1. Air Force Research Laboratory. Integrated cockpit sensing. 2022. [Accessed Sept. 5, 2023]. Available from https://afresearchlab.com/wp-content/uploads/2022/03/ICS_ProjOverview_Booklet_Distro-A-2022-0438.pdf.
2. Akiba T, Sano S, Yanase T, Ohta T, Koyama M. Optuna: a next-generation hyperparameter optimization framework. Proceedings of the 25th ACM SIGKDD; July 2019; International conference on knowledge discovery & data mining. New York: ACM; 2019:2623-2631.
3. Cao XS, Wang YC, Xu L, Yang CB, Wang B, et al. Visual symptoms and G-induced loss of consciousness in 594 Chinese Air Force aircrew—a questionnaire survey. *Mil Med.* 2012; 177(2):163–168.
4. Green ND, Ford SA. G-induced loss of consciousness: retrospective survey results from 2259 military aircrew. *Aviat Space Environ Med.* 2006; 77(6):619–623.
5. Kurihara K, Kikukawa A, Kobayashi A, Nakadate T. Frontal cortical oxygenation changes during gravity-induced loss of consciousness in humans: a near-infrared spatially resolved spectroscopic study. *J Appl Physiol.* 2007; 103(4):1326–1331.
6. Lockheed Martin Corporation. Automatic Ground Collision Avoidance System (Auto GCAS). 2023. [Accessed Sept. 5, 2023]. Available from <https://www.lockheedmartin.com/en-us/products/autogcas.html>.
7. Lyons TJ, Kraft NO, Copley GB, Davenport C, Grayson K, et al. Analysis of mission and aircraft factors in G-induced loss of consciousness in the USAF: 1982–2002. *Aviat Space Environ Med.* 2004; 75(6):479–482.
8. Newman DG. Characteristics of a G-LOC episode. In: Newman DG. High G flight: physiological effects and countermeasures. Farnham, UK: Ashgate Publishing Ltd.; 2015:65–68.
9. Newman DG. Individual factors affecting G tolerance. In: Newman DG. High G flight: physiological effects and countermeasures. Farnham, UK: Ashgate Publishing Ltd.; 2015:118–119.
10. Ohruai N, Fujita M, Kikukawa A, Kuramoto K, Kobayashi A, et al. Countermeasure and monitoring for gravity-induced loss of consciousness (G-LOC). *Aeromedical Laboratory Reports.* 2016; 56(3):43–66 [in Japanese].
11. Ohruai N, Fujita M, Kuramoto K, Kikukawa A, Kobayashi A, et al. Preliminary trial to establish an abbreviated formula for gravity-induced loss of consciousness (G-LOC) prediction. *Aeromedical Laboratory Reports.* 2017; 57(2):15–26 [in Japanese].
12. Ohruai N, Fujita M, Tsuruhara A, Kuramoto K, Kikukawa A, et al. Cerebral oxyhemoglobin concentration changes within five seconds after the onset of high G_z. *Aeromedical Laboratory Reports.* 2019; 59(1):1–12 [in Japanese].
13. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion, B, et al. Scikit-learn: machine learning in Python. *Journal of Machine Learning Research.* 2011; 12:2825–2830.
14. Rickards CA, Newman DG. G-induced visual and cognitive disturbances in a survey of 65 operational fighter pilots. *Aviat Space Environ Med.* 2005; 76(5):496–500.
15. Van Rossum G, Drake FL. Python 3 reference manual. Scotts Valley (CA): CreateSpace; 2009:1–242.
16. Whinnery T, Forster EM. The +G_z-induced loss of consciousness curve. *Extrem Physiol Med.* 2013; 2(1):19.