scientific reports

OPEN

A new open-access platform for measuring and sharing mTBI data

August G. Domel[1,12], Samuel J. Raymond[1,12✉], Chiara Giordano[1,12], Yuzhe Liu[1], Seyed Abdolmajid Yousefsani[1], Michael Fanton[2], Nicholas J. Cecchi[1], Olga Vovk[3], Ileana Pirozzi[1], Ali Kight[1], Brett Avery[4], Athanasia Boumis[4], Tyler Fetters[3], Simran Jandu[4], William M. Mehring[4], Sam Monga[5,6], Nicole Mouchawar[7], India Rangel[4], Eli Rice[4], Pritha Roy[4], Sohrab Sami[4], Heer Singh[4], Lyndia Wu[1,8], Calvin Kuo[2,9], Michael Zeineh[7], Gerald Grant[10,11] & David B. Camarillo[1,2,11]

Despite numerous research efforts, the precise mechanisms of concussion have yet to be fully uncovered. Clinical studies on high-risk populations, such as contact sports athletes, have become more common and give insight on the link between impact severity and brain injury risk through the use of wearable sensors and neurological testing. However, as the number of institutions operating these studies grows, there is a growing need for a platform to share these data to facilitate our understanding of concussion mechanisms and aid in the development of suitable diagnostic tools. To that end, this paper puts forth two contributions: (1) a centralized, open-access platform for storing and sharing head impact data, in collaboration with the Federal Interagency Traumatic Brain Injury Research informatics system (FITBIR), and (2) a deep learning impact detection algorithm (MiGNet) to differentiate between true head impacts and false positives for the previously biomechanically validated instrumented mouthguard sensor (MiG2.0), all of which easily interfaces with FITBIR. We report 96% accuracy using MiGNet, based on a neural network model, improving on previous work based on Support Vector Machines achieving 91% accuracy, on an out of sample dataset of high school and collegiate football head impacts. The integrated MiG2.0 and FITBIR system serve as a collaborative research tool to be disseminated across multiple institutions towards creating a standardized dataset for furthering the knowledge of concussion biomechanics.

In this study, we seek to address two major outstanding problems that we believe exist with pinning down the underlying biomechanical mechanisms of brain injury, and more specifically concussion. Because of the rarity of measuring biomechanical data during a concussion, we believe the first of two major issues is the lack of a large, high-quality head impact dataset with associated clinical outcomes. Secondly, because many of the systems, for example the Head Impact Telemetry System (HITS), used to capture the biomechanical data of head impacts often generate many false impacts, this can further generate confusing and conflicting results among studies. The only way around this is to video verify and confirm every real and fake impact measured, which is a tedious process that hampers and slows research significantly, or as is often the case, to publish the data as if all impacts are real. This is something that has hampered previous efforts at a large head impact dataset, such as that of the Concussion Assessment, Research and Education (CARE) Consortium. Given these two major issues, the goal of this study is two-fold: (1) to put forth a much needed centralized and open-source platform

[1]Department of Bioengineering, Stanford University, Stanford, CA 94305, USA. [2]Department of Mechanical Engineering, Stanford University, Stanford, CA 94305, USA. [3]General Dynamics Information Technology, 3150 Fairview Park Drive, Falls Church, VA 22042, USA. [4]Stanford Center for Clinical Research, Stanford University, Stanford, CA 94305, USA. [5]Intel Sports, Intel Corporation, Santa Clara, CA 95054, USA. [6]Asterix Inc, Toronto, ON M4B 1B3, Canada. [7]Department of Radiology, Stanford University, 300 Pasteur Dr, MC 5105, Stanford, CA, USA. [8]Department of Mechanical Engineering, University of British Columbia, 6250 Applied Science Lane, Vancouver, BC V6T 1Z4, Canada. [9]Department of Biomedical Engineering, University of British Columbia, Lower Mall Research Station, Vancouver, BC 2259V6T1Z4, Canada. [10]Department of Neurology, Stanford University, Stanford, CA, 94305, USA. [11]Department of Neurosurgery, Stanford University, Stanford, CA 94305, USA. [12]These authors contributed equally: August G. Domel, Samuel J. Raymond and Chiara Giordano. ✉email: sjray@stanford.edu

for storing and sharing in vivo head impact data and (2) to put forth a deep learning impact detection algorithm for differentiating between true and fake impacts of the extensively validated (i.e. high-quality) instrumented mouthguard sensor which has been designed to easily interface with the open-source platform. We seek to provide the research community with this integrated platform for collecting, processing, and sharing high-quality, standardized, biomechanical head impact data, to create a comprehensive injury and non-injury dataset. The proposed research platform is developed in collaboration with and hosted by the National Institutes of Health's (NIH) Federal Interagency Traumatic Brain Injury Research (FITBIR) informatics system[1]. Our instrumented mouthguard device collects kinematic data according to the recommended Common Data Elements (CDEs) from the National Institute of Neurological Disorders and Stroke (NINDS), which allows integration and sharing of data collected among different investigators via the platform[2]. Furthermore, our NIH R24 grant allows for the funding of 20 instrumented mouthguards for each site that joins the study to contribute to this platform.

An open-source platform, such as this, is of such great importance because of the known dangers associated with brain injury. Traumatic Brain Injury (TBI) is a leading cause of death and disability; however, more recently it has been proposed that even mild TBI (mTBI), or concussion, can increase the risk for neurodegenerative diseases, such as Alzheimer's[3,4]. Researchers have accumulated evidence suggesting that even subconcussive impacts, previously thought to be innocuous, result in lasting neurophysiological changes[5–7]. Therefore, there is a pressing need to uncover the link between the number and severity of head impacts sustained in a lifespan, and how this correlates to acute symptomatology and affects risk of long-term chronic neurodegeneration. The scientific community would greatly benefit from a metric for concussion diagnosis, such as a potential concussion threshold. However, in order to do this, accurate biomechanical measurements of head impacts is crucial.

Previous efforts to measure head impact kinematics have leveraged sensor devices mounted to athlete's heads through headbands, ear plugs, helmets, bite bars, skin patches, and mouthguards[8–12]. However, the lack of robust coupling to the skull of some of these devices has produced inaccuracies in the measurements observed both in laboratory and field settings[11,13–16]. Thus, it is clear that when using them with athletes in field studies, the conclusions drawn have limitations[12]. When this issue is coupled with the lack of impact detection algorithms or video verification of impacts used by many studies, many fake impacts will confuse the conclusions, and the results compared between studies can be conflicting[17–20]. Some studies have found no correlation between biomechanical metrics and the number of incurred events, severity, or time to symptoms[20,21]. However, others report injury thresholds based on head accelerations[22,23], which are ultimately used to rate how safe certain protective systems are that are on the market. Thus, ongoing research aimed at establishing the relationship between head impact biomechanics and brain health is hampered by the lack of a large, high-quality, reliable concussive and subconcussive dataset measured in 6 degrees of freedom[20–23].

In recent years, it has become clear that mouth wearables (i.e. instrumented mouthguards) have the tightest coupling to the skull of current head kinematic wearables[9]. Specifically, the Stanford instrumented mouthguard (herein referred to as the "MiG2.0") has been very recently extensively validated to ensure high-quality data is collected for head impacts[24]. Using this instrumented mouthguard, along with a newly developed deep learning algorithm to remove fake impacts, we believe we can produce high quality data to lead to more meaningful study conclusions. Recently, a number of studies have employed machine learning approaches to remove fake impacts[23–27]. Wu et al. used a support vector machine (SVM) to detect head impacts in American football players[27]. Others, such as the work conducted in[28] by Motiwale et al. have used deep learning to analyze biomechanical signals. A benefit of using deep learning tools, like neural networks, is that the features of interest need not be chosen by the researchers but instead are borne out of the data during the training process. Adding more data to the system typically improves the model more and more for these kinds of algorithms. This makes the expansion of the database ever more useful. An interesting comparison of the predictive power between SVM and deep neural networks is the choice of features. In the SVM these need to be engineered, known apriori, so that the model can use them to learn the data. However, in neural networks, these features are discovered during the training process itself. The benefit of this is that the features are typically optimally chosen to explain the data, the downside being that the knowledge of what these features relate to can be difficult, if not impossible, to explain. This tradeoff between accuracy and explainability has not yet been explored in the detection of head impacts, as it relates to mTBI, and is focused on in this work. While the accuracy of detecting head impacts is certainly the most important metric, understanding what is sacrificed to achieve this is also valuable in terms of model selection. In this study, we present a deep learning model to analyze and remove fake impacts and compare it to an SVM to understand its benefits.

With the extensively validated MiG2.0 and its newly improved deep learning impact detection algorithms discussed in this paper, we hope for this manuscript to be a call to collaborators all over to join in the study and contribute to the new open-source platform. We hope that this platform and associated high-quality data collection will more quickly and robustly shed light on the mechanisms of brain injury and concussion.

Materials and methods
The MiG. 2.0.   We have developed an instrumented mouthguard, the MiG2.0, for measuring linear and angular head kinematics during an impact. Several images of the molded mouthguard as well as the electronics embedded within can be found in Fig. 1. The device is capable of sensing six degrees of freedom (6DOF) kinematics via a triaxial accelerometer (H3LIS331DL, ST Microelectronics, Geneva, Switzerland) with maximum range of ± 400 g, and a triaxial gyroscope (ITG-3701, InvenSense Inc., San Jose, CA, US) with maximum range of ± 4000 degrees/sec. Linear acceleration is acquired at 1000 Hz, while angular velocity is acquired at 8000 Hz[17]. Data is temporarily stored in an on-board buffer, which allows the acquisition of pre- and post-trigger data associated with an event. When the linear acceleration threshold value set by the user is crossed at any of the linear acceleration components, an event is written to the on-board flash memory chip. Trigger threshold, pre-trigger
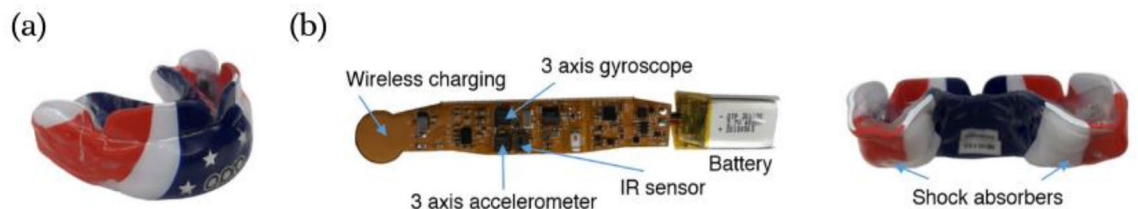
Figure 1. MiG2.0 design. (**a**) The Stanford instrumented mouthguard is a custom-made mouthguard instrumented with a triaxial accelerometer and a triaxial gyroscope for measurements of head kinematics. (**b**) The sensory board is located at the incisors. For in mouth sensing the MiG2.0 is equipped with an infrared sensor. Unique shock absorbers are placed at the molars to reduce external disturbances.

time, event duration, and sample rate of the inertial measurement units (IMUs) are user adjustable. The device is also capable of detecting when it is being worn when using a proximity sensing via an infrared emitter-receiver IR pair (TMD2771, Texas Advanced Optoelectronic Solutions Inc., Plano, TX, US) which detects light reflected off of the teeth. The sensor board is completely sealed between three layers of ethylene vinyl acetate (EVA) material and communication occurs via Bluetooth. A tight fit to the dentition is achieved by forming the EVA material around a player-specific dental model. IMUs are located at the incisors to isolate sensors from external disturbances. Furthermore, unique shock absorbers are placed at the molars to reduce mandible disturbances occurring in chewing, biting, and talking with the device on the teeth, as discussed by Kuo et al.[15]. It is important to note that the accuracy of all kinematic data for the MiG2.0 has been previously validated and reported in a recent manuscript[24].

Data download using BiteMac and upload to shared FITBIR platform. To access the kinematic data collected by the mouthguard, we developed an application, BiteMac, to interface with the MiG2.0. BiteMac is a free iOS and macOS application that enables quick, wireless, and intuitive download of kinematic data in comma separated value (CSV) format using the recommended common data elements (CDE) from the National Institute of Neurological Disorders and Strokes (NINDS) for collecting sensor data. CDEs were developed by a group of experts to facilitate collaboration across investigators, as well as interconnectivity with other informatics platforms[29]. We have worked with the Federal Interagency Traumatic Brain Injury Research (FITBIR) agency to create an open-access platform on their website to share standardized concussion data collected using the MiG2.0 or any other head impact kinematic sensor (Methods 2.3).

Finally, we have developed a website which will enable researchers to register to use the MiG2.0 for their studies[2]. This website also contains the necessary scripts to process and standardize data output from the MiG2.0, as well as perform impact classification using our machine learning algorithm described later.

The FITBIR online platform. The FITBIR informatics system[30] was developed to share data (whether taken with the MiG2.0 or another similar high-quality device) across the TBI research field, in order to facilitate collaboration between laboratories and enhance interconnectivity with other informatics platforms thus allowing for aggregation, integration, and rigorous comparison of data[27]. Our integrated system provides the common data definitions and standards that are required for community-wide sharing, and qualified researchers can join the study and request access privileges to the data stored in FITBIR[30]. The FITBIR data dictionary incorporates and extends the CDE definitions developed by the NINDS CDE Workgroup, whereby domestic and international neuroscience experts developed disease-specific CDEs[31,32]. To date, CDEs have been developed for head kinematics sensors, a full list of which can be found at[1], the SCAT-5 (Sport Concussion Assessment Tool), and the ImPACT exam (Immediate Post-Concussion Assessment Testing)[1,32]. For example, these CDEs include information on the participants as well as their impact data, ranging from the participants' weight and heights to the data measured by the instrumented mouthguards and via the sideline neurocognitive tests. Data can be uploaded to our study in the FITBIR platform in CSV format, and we provide collaborating investigators with scripts to harmonize the MiG2.0 data with the FITBIR data input format. Data upload can be performed on the FITBIR website (the overall workflow for this is illustrated in Fig. 2) and should be carried out at the end of data collection by all participating investigators, once per year. Note that data is anonymized to care for all protected health information (PHI). As previously discussed in "Data download using BiteMac and upload to shared FITBIR platform", researchers are able to register to join the study via our website[2], and then contribute to the data repository on FITBIR upon conducting their study.

MiG2.0 deployment to generate impact detection training and testing dataset. We deployed the MiG2.0 during practices and games to one college football team, Stanford University, and three local Bay Area high schools. We video analyzed the data from the collegiate players (n = 12), including 5 outside linebackers, 1 inside linebacker, 1 running back, 2 wide receivers, 1 defensive end, 1 offensive tackle, and 1 center, over the course of 17 practice and game days. We also video analyzed the data from the high school players (n = 49) over the course of 17 practice and game days. The high schoolers typically play multiple positions, unlike collegiate players, so our data sample included almost all football positions, except a kicker. The collegiate football games followed the National Collegiate Athletic Association (NCAA) Football Rules, and the practices were typically 2 h in duration. The high school games and practices were also approximately 2 h in duration, in
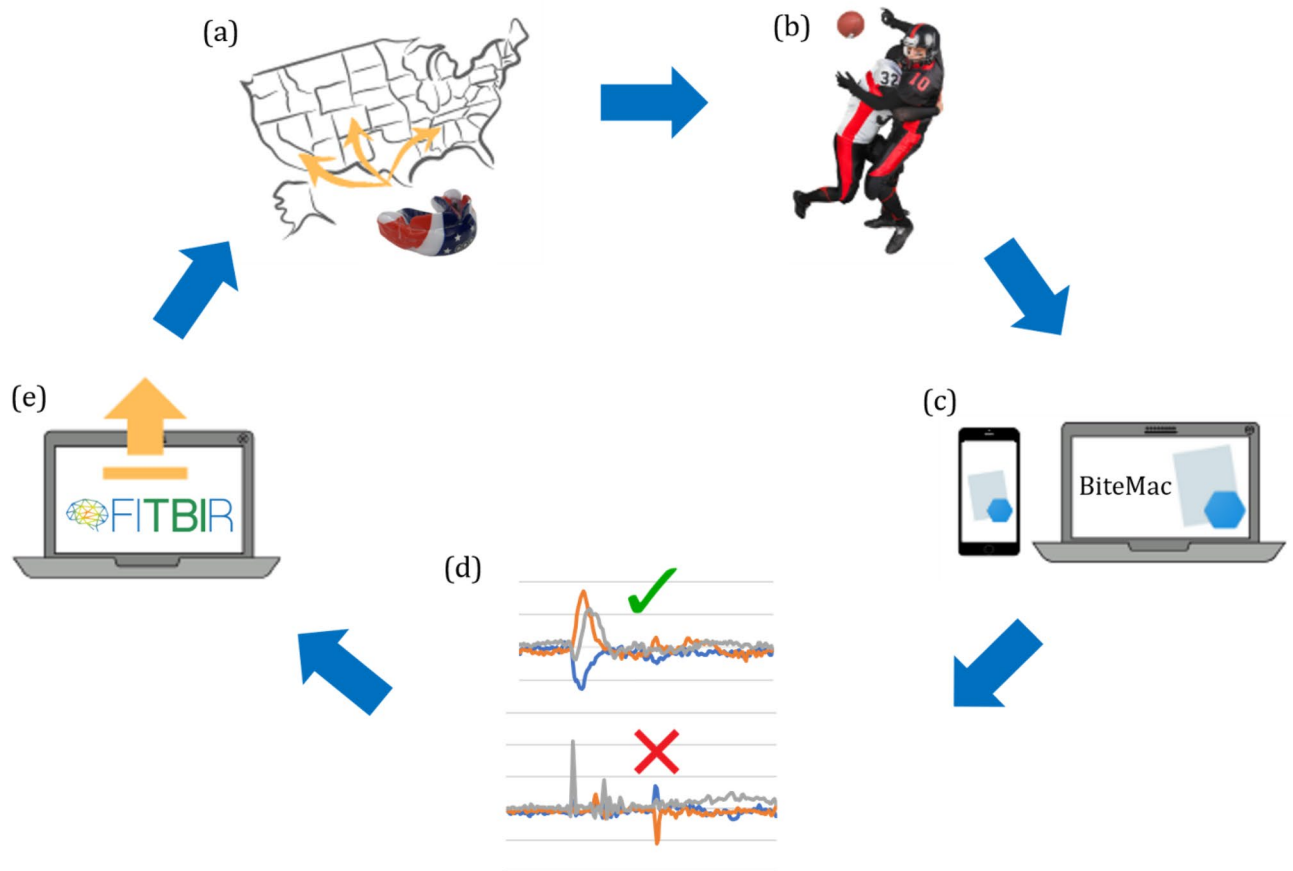
Figure 2. Visual representation of the proposed integrated platform for collection, processing and sharing of mTBI data. (**a**) Distribution of MiG2.0 to partner investigators who enroll in the NIH funded study (in collaboration with FITBIR) to generate an open-access platform for sharing standardized head kinematic and concussion data; (**b**) MiG2.0′s are deployed to collect field data; (**c**) Raw data can be accessed and downloaded from the mouthguards directly by the investigators using the custom-made BiteMac application; (**d**) The raw data will be processed with our impact detection algorithm to distinguish true impacts from fake impacts; (**e**) Uploading to FITBIR platform can be carried out by investigators to share mTBI data and encourage other investigators to join the study as well.

accordance with the California Interscholastic Federation (CIF). The human subject protocols were approved by the Stanford Administrative Panels for the Protection of Human Subjects (IRB-45932 and IRB-46537). We conducted data collection and performed video analysis in accordance with the institutional review boards' guidelines and regulations. Prior to deploying the MiG2.0, research staff obtained informed consent from all participants. Participants over the age of 18 provided their own informed consent. Participants under the age of 18 had a legally authorized representatives provide consent and then provided their own assent. All consents and assents were written and signed by participants or legally authorized representatives. The informed consent process used here followed the Stanford Institutional Review Board guidelines.

The obtained kinematics data was used to train and validate an impact classifier in part based on the method published by Wu and coworkers for detecting field football head impacts[31,32]. Retraining the classifier was important because the MiG2.0 does not have the same mechanical and electrical design as previous iterations of the device[33]. In our dataset, we included video-verified head impacts, while video-verified noncontact recordings were defined as noncontact events. To assign the mouthguard kinematics to one of the aforementioned classes, we synchronized mouthguard event timestamps to timestamps from practice and game videos, and all events (as detected by over 10 g linear acceleration in any direction) that were on camera were video verified as either a real impact or noncontact event by members of the research team. Only data recordings where contact (or noncontact) could be clearly and confidently identified were included in the training/validation dataset. Videos were acquired from at minimum 2 angles per field at 30 frames per second. However, as over roughly 70% of these potential impacts occurred out of view for the cameras, an improvement to the tracking of players would yield a larger dataset. In the future, we are hoping to track individual players throughout the entire practice or game in order to capture many more impacts and to get a better sense of the impacts that are off camera.

MiGNet: neural network classifier for impact detection validation on field impacts. In this section, we present a deep learning approach to automatically extract important features to distinguish between real and false impacts to a higher degree of accuracy using mouthguard time series data as the input. Deep learning

has become common in many areas of biomedical/biomechanical research[34–36] where the signals generated are complex and a more automated feature extraction process is desired. One trade off to this approach, compared with previous work, such as the SVM classifier used in[33] is that the features are no longer selected or directly related to any engineered metric. The SVM model used in[33] is compared to this new deep learning model for our new dataset with the advantages and disadvantages compared.

The dataset of mouthguard signals. Our dataset consisted of a collection of 1×200 timeseries signals. These signals are produced by the mouth guards presented in this and previous work. When an impact force exceeds a threshold of 10 g the onboard recorder stores the previous 50 ms and following 150 ms of kinematic data (see "MiG2.0 deployment to generate impact detection training and testing dataset"). This kinematic data is the linear acceleration in 3 dimensions and angular velocities in 3 dimensions. Once a mouth guard is returned for processing, a manual tagging process is conducted whereby video analysis is used to confirm if a recorded impact corresponds to a video verified head impact from the video recording of the player. This creates a labeled dataset of false impact readings and true impact readings. These two classes are then used to train the impact detection algorithm in an attempt to remove the need for the manual, and slow, process of video verification.

The dataset used to train this new SVM and the new MiGNet neural network consisted of 358 true impacts and 500 false impacts. These video-verified impact signals were obtained from the MiG2.0 mouthguards worn by the athletes. The events were classified as true and false impacts by video-confirmation, as described in previous sections. For the support vector dimensions, a feature extraction process was conducted for the data with a column selection process from[28] used to identify the most relevant features.

Augmentation and Imbalanced Data Strategy for MiGNet Training. Due to the complexity of the signals and the imbalance in the data, two strategies were executed to achieve the best performance for the network training. Firstly, to increase the number of signals used to train with, an augmentation of the signals was performed whereby the signal of a kinematic variable was shifted along the time axis by a unit of milliseconds. The total number of augmented signals from one original sample was a controlled variable. To ensure the signals were still representative of realistic impacts, and to ensure sufficient data for training, 5 augmented samples were obtained, meaning that the signal was offset by 1-5 ms for each collected signal. However, it is important to note that the test set used for later evaluation was unaltered and not used in training.

In addition to the modest dataset size that was available, this dataset was also imbalanced. This resulted in suboptimal performance of the network originally. By calculating the different weighting for each class such that their importance is equally treated in the loss function used during training, this imbalance can be treated more fairly. This essentially allows for the loss-function to pay more attention to losses for the true events, and less so for the false events. More details of treating imbalanced classes can be found in[37].

As the input data was in a 1D time-varying format, the network architectures were chosen accordingly. Several variations of network layer size and architecture were iterated until the final choice of network, shown in the schematic of Fig. 3. This included a series of 1D convolution layers that were later combined into a 2D output where 2D convolutions were then performed, until culminating in a Softmax output layer. 1D convolutions were used in preference to Long Short-term Memory (LSTM) due to the efficiency of 1D convolution layers[38,39]. Stochastic gradient descent was used to train the network using the Keras framework with Google's TensorFlow as the background engine. Batch sizes of 32 were iterated for 20 epochs with a learning rate of 0.01 and momentum of 0.9.

Finally, the comparative performance of the neural net and the SVM was evaluated using sensitivity, specificity, accuracy and precision as described below:

$$Sensitivity = \frac{TP}{TP + FN} \quad Specificity = \frac{TN}{TN + FP}$$

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad Precision = \frac{TP}{TP + FP}$$

True positive (TP) was an impact classified as impact, true negative (TN) was a nonimpact classified as a nonimpact, a false positive (FP) was a nonimpact classified as an impact, and a false negative (FN) was an impact classified as a nonimpact.

Results

MiGNet architecture determination.
The MiGNet architecture was chosen by comparing a number of different layer configurations. As the input signals are all designed to be of a constant size, a 1D convolutional layer was always used and an LSTM approach was not investigated. Several different models are listed below and the results are shown in Table 1. While most of these models have only a modest depth after testing several deep 1D convolutional layers we found minimal difference in performance. Similarly, the number of 2D convolutional layers did not significantly affect performance metrics. For the final iteration of the MiGNet, two 1D convolutional layers and one 2D convolutional layer was used in order to increase the computational speed of the net by providing fewer parameters. For the last hidden layer, Global Average Pooling resulted in the best accuracy due to its emphasis on connection between features, maps, and categories. The small size and therefore minimal computing power needed is also a benefit as these models are intended to be placed on the mouthguard hardware in the future with a minimal computational footprint desired. The different model architectures that were compared are listed below:

5
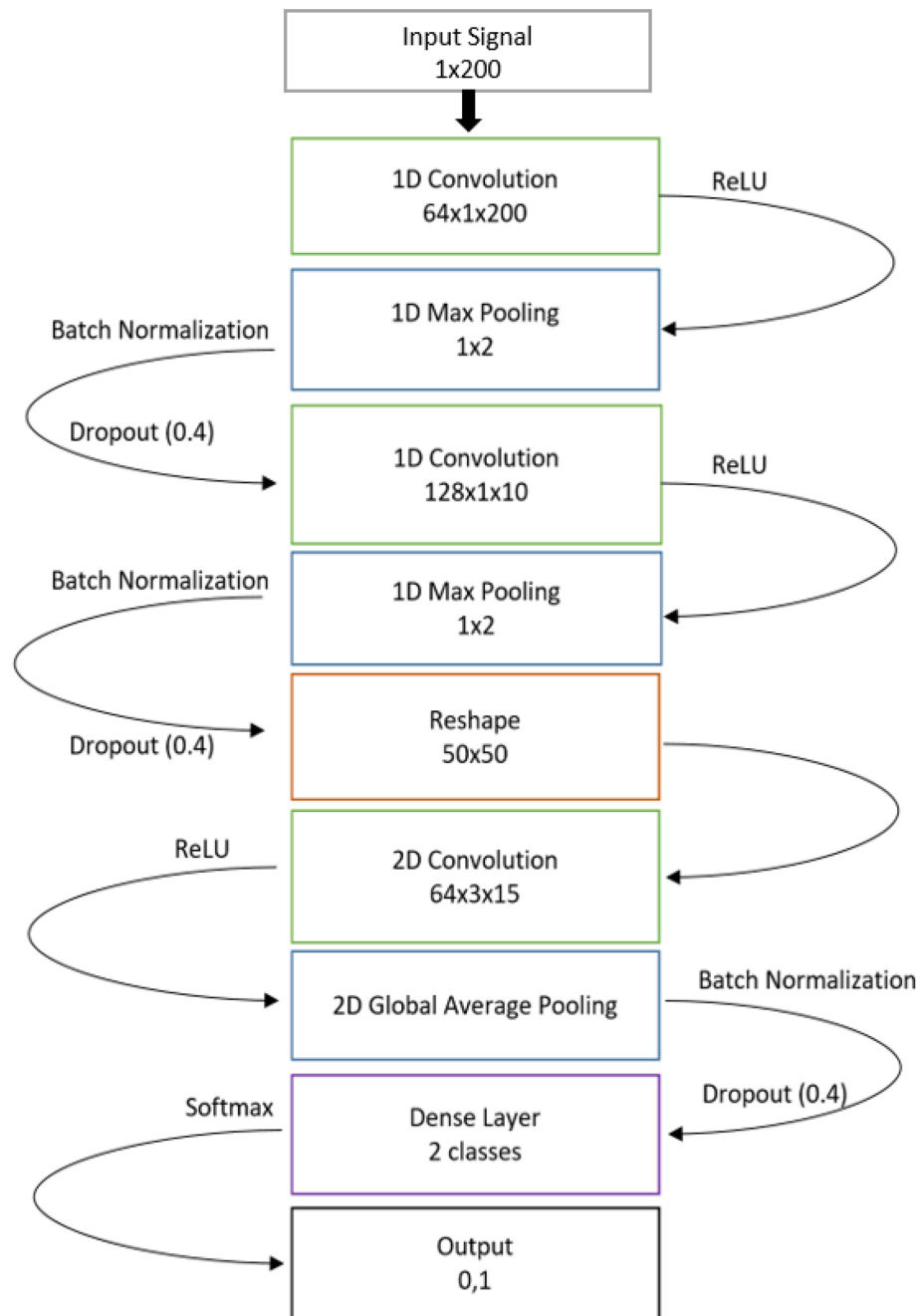
Figure 3. MiGNet architecture schematic. The 1D convolutional layers act to extract high-level features of the motion signal, feeding into a 2D convolution which fuses the sensor signals together.

	TP	FP	TN	FN	Sensitivity	Specificity	Accuracy	PPV	NPV
Model 1	35	5	410	15	70%	99%	96%	0.88	0.96
Model 2	36	5	410	14	72%	99%	96%	0.88	0.97
Model 3	32	16	399	18	64%	96%	93%	0.67	0.96
Model 4	38	6	409	12	76%	99%	96%	0.86	0.97

Table 1. MiGNet architecture performances. Performance measures for the different trained MiGNet architectures.

	Test 1—SVM	Test 1—MiGNet	Test 2—MiGNet
Dataset size	165	165	512
Sensitivity	86%	97%	76%
Specificity	94%	90%	99%
Accuracy	91%	93%	96%
Precision	90%	86%	86%

Table 2. MiGNet and SVM performance. Performance measures for the trained MiGNet and SVM showing the performance on the initial dataset (Test 1) and the larger dataset used to improve the MiGNet (Test 2).

Model 1: 2x [1D convolution layers (ReLU and 64 filters, filter size 3×1), a max pooling, 2 more 1D convolution layers (ReLU and 128 filters, filter size 3×1)], Global average pooling, a dropout layer with 50%, output layer with sigmoid activation.

Model 2: 3x [1D convolution layer (128 filters, 1×8 (for the first layer)), then 1×5 (for the second layer), then 1×3 (for the third layer) filter sizes, ReLU activation), batch normalization], global average pooling, output layer with Softmax activation.

Model 3: 2x [1D convolution layer (48 (for the first) and 96 (for the second) filters, 15×1 filter size, ReLU activation), max pooling (pool = 2), dropout layer 40%], a reshaped 2D combination of layer one and layer 2 to create a 2D matrix of combined layers, 2D convolution (96 filters, 3×15 patch, 3×1 strides, ReLU activation), global average pooling, dropout layer with 40%, output layer with sigmoid activation.

Model 4: 2x [1D convolution layer (64 (for the first) and 128 (for the second) filters, 5×1 (for the first) and 10×1 (for the second) filter sizes, ReLU activation), max pooling (pool = 2), batch normalization layer, dropout layer 40%], a reshaped 2D combination of layer one and layer 2 to create a 2D matrix of combined layers, 2D convolution (64 filters, 3×15 patch, 3×1 stride size, ReLU activation), global average pooling, batch normalization layer, dropout layer with 40%, output layer with Softmax activation.

These models were tested on the same data and, with the overall goal of minimizing false negative being more important than minimizing false positives, model 4 was chosen for the rest of this work. As more data is collected in this study, refinement of this architecture will be conducted as needed. With the overall goal of minimizing false negative more important than minimizing false positives, Model 4 was chosen for the rest of this work. As more data is collected in this study, refinement of this architecture will be conducted as needed. The final structure of the model used is depicted in Fig. 3.

MiGNet vs. SVM impact detection validation. The SVM features were extracted using the inbuilt MATLAB function, *sequentialfs*, and the SVM itself was trained using the inbuilt MATLAB function, *fitcsvm*. Testing these two models on an out of sample set of 65 impacts and 100 non impacts yielded the results shown in Table 2. For this dataset, the two methods are comparable in the four metrics, with the MiGNet slightly outperforming the SVM overall. However, with more data, these machine learning models typically improve performance. Unlike the neural network, though, for the SVM a larger dataset requires that the feature selection process be repeated. As this feature selection process involves a matrix decomposition for each feature, this process is incredibly expensive and largely infeasible where other options are available. For the MiGNet, adding more data to train does still have the effect of increasing the training time, but at a much more efficient rate. In an attempt to improve the MiGNet performance, a larger dataset was applied to the training, with 3730 nonimpact and true impact signals. Retraining on this larger dataset for the SVM would take an infeasible amount of time; however, the MiGNet model was retrained within the time it took to train the smaller SVM. With this extra data, the MiGNet's final performance metrics are shown in Table 2 with the MiGNet reporting 96% accuracy, 86% precision, 99% specificity, and 76% sensitivity on the out of sample test set.

Discussion
Shared FITBIR Platform. The purpose of this work was to propose an open-source platform for collecting, analyzing, and sharing head impact data, including head impact kinematics and the resulting health outcomes, such as sideline neurocognitive test results and whether a concussion occurred. This proposed platform would help investigators to further study links between head impact kinematics and various proposed injury metrics, thus giving further insights into the biomechanical mechanisms of concussive and subconcussive impacts. Furthermore, in this manuscript, we described a wearable device to accurately detect and capture the features of high-speed head impacts and a pipeline to easily integrate the data captured with this device into the FITBIR platform. The discussed device has been distributed to a selected group of investigators currently enrolled in an NIH funded project aimed at elucidating the biomechanical underpinnings of concussion. The standardization of a high-quality concussion dataset to establish a link between head impact kinematics and brain injury is paramount especially since these injuries are rare for any one individual. Given this, it appears that the necessary data can be more quickly and effectively gathered through large-scale collaborative efforts involving multiple institutions through the dissemination of a validated head impact measurement device. The MiG2.0 collects data in a standardized form employing the NINDS recommended CDE, thus facilitating collaboration and interconnectivity with informatics platforms such as FITBIR, where data will be released in an open-source format at the time of project completion in the next couple of years. We believe that the collective knowledge gained from this

collaborative effort will improve our understanding of concussion and aid in determining the now elusive link between head kinematics and brain injury.

Evaluation of the MiGNet Impact Detection Algorithm. Combining the use of head impact sensor data with independent video analysis, we trained and evaluated an impact detection classifier: a convolutional neural net trained and tested with high school and collegiate field data. The MiGNet architecture was developed using a greedy optimization scheme, and a parameter sweep was used to find the optimal filter size, kernel width and dropout percent.

As a deep learning product, the MiGNet benefits from a large availability of data, and while the augmentation strategy described earlier allows for sufficient training data, additional real-world data will improve this platform's performance. As the proposed integrated platform becomes widespread amongst the research community, the head impact dataset will enable more accurate training of the neural net, leading to enhanced impact detection through MiGNet workflow. Though it may take years to obtain a sizable dataset, the performance of the MiGNet gives promise that high-accuracy impact detection may be further achieved based on this methodology.

Limitations. The classifier described in this study was only evaluated for use with the MiG2.0 in American football. Retraining of the classifier is potentially necessary for any device that does not have the same mechanical and electrical design as the MiG2.0. We also recognize the need to test impact detection in other contact sports, both helmeted and unhelmeted to determine whether retraining the algorithm is a necessity given that other sports will likely involve forms of contact and/or player behaviors that may not have been analyzed by the current classifier. Moreover, using video information to establish ground truth data labels has some limitations that could perhaps be addressed by advances in computer vision. In this study, we rigorously checked all video labels to ensure high confidence in the training set and only used data that was confidently and clearly verified on video. Almost all mouthguard data recordings from players buried in a pile were excluded, although helmet contact is likely to occur in such situations and may be of a unique nature, which we have disregarded. Thus, we have limited our training to only very visible events. We would also like to note that we don't consider potential impacts that were missed by the mouthguard, which tend to be because they are below the minimum threshold set to trigger an impact, we instead focus only on the mouthguard recorded events for this study.

Conclusion

There are many unanswered questions underlying the biomechanics of brain trauma: we do not know what force levels could lead to a concussive event or whether this threshold changes with age, gender, or type of activity. For the research field to answer all of these questions, an integrated platform for collecting, post-processing, and sharing standardized, high-quality head impact data is essential. In previous work, we demonstrated that the MiG2.0 can collect high-quality head impact kinematics in helmeted contact sports like American Football. In the current study, we provided a validated machine learning algorithm for impact detection and classification of on-field events. Overall, by collecting standardized common data elements, the MiG2.0 (and MiGNet) and FITBIR platform facilitate collaboration across investigators. This feature makes the MiG2.0 and FITBIR platform appropriate for enabling and fostering across-borders and multi-institutional collaboration efforts to produce a standardized, high-quality database for head impacts, including the head impact kinematics and resulting health outcomes, such as whether concussion occurred. We are hopeful that the knowledge gained from this dissemination effort will improve our basic understanding of the science of concussion.

References

1. NINDS 2020, *CDE Catalog. Common Data Elements*. https://www.commondataelements.ninds.nih.gov/cde-catalog (2020).
2. Camarillo, D. *Stanford Instrumented Mouthguard* (Department of Bioengineering, Stanford University). http://mouthguards.stanford.edu/login (2020).
3. Zunt, J. R. *et al.* Global, regional, and national burden of traumatic brain injury and spinal cord injury, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol.* **17**, 1061–1082 (2018).
4. Kamins, J. & Giza, C. C. Concussion—Mild traumatic brain injury: recoverable injury with potential for serious sequelae. *Neurosurg. Clin.* **27**(4), 441–452. https://doi.org/10.1016/j.nec.2016.05.005 (2016).
5. Guskiewicz, K. M. *et al.* Association between recurrent concussion and late-life cognitive impairment in retired professional football players. *Neurosurgery* **57**, 719–726 (2005).
6. Stemper, B. D. *et al.* Comparison of head impact exposure between concussed football 522 athletes and matched controls: Evidence for a possible second mechanism of sport523 related concussion. *Ann. Biomed. Eng.* https://doi.org/10.1007/s10439-018-02136-6 (2018).
7. Hernandez, F. *et al.* Lateral impacts correlate with falx cerebri displacement and corpus callosum trauma in sports-related concussions. *Biomech. Model Mechanobiol.* **18**, 631–649. https://doi.org/10.1007/s10237-018-01106-0 (2019).
8. Bartsch, A., Samorezov, S., Benzel, E., Miele, V. & Brett, D. Validation of an "intelligent mouthguard" single event head impact dosimeter (No. 2014-22-0001). SAE Technical Paper (2014).
9. Camarillo, D. B., Shull, P. B., Mattson, J., Shultz, R. & Garza, D. An instrumented mouthguard for measuring linear and angular head impact kinematics in american football. *Ann. Biomed. Eng.* **41**, 1939–1949 (2013).
10. Rowson, S. *et al.* A six degree of freedom head acceleration measurement device for use in football. *J. Appl. Biomech.* **27**, 8–14 (2011).
11. Wu, L. C. *et al.* In vivo evaluation of wearable head impact sensors. *Ann. Biomed. Eng.* **44**(4), 1234–1245 (2016).
12. King, D. A., Hume, P. A., Gissane, C. & Clark, T. N. Similar head impact acceleration measured using instrumented ear patches in a junior rugby union team during matches in comparison with other sports. *J. Neurosurg. Pediatr.* **18**(1), 65–72 (2016).

13. Patton, D. A. A review of instrumented equipment to investigate head impacts in sport. *Appl. Bionics Biomech.* https://doi.org/10.1155/2016/7049743 (2016).
14. Jadischke, R., Viano, D. C., Dau, N., King, A. I. & McCarthy, J. On the accuracy of the head impact telemetry (hit) system used in football helmets. *J. Biomech.* **46**, 2310–2315 (2013).
15. Kuo, C. *et al.* Effect of the mandible on mouthguard measurements of head kinematics. *J. Biomech.* **49**, 1845–1853 (2016).
16. Siegmund, G. P., Guskiewicz, K. M., Marshall, S. W., DeMarco, A. L. & Bonin, S. J. Laboratory validation of two wearable sensor systems for measuring head impact severity in football players. *Ann. Biomed. Eng.* **44**, 1257–1274 (2016).
17. Wu, L. C. *et al.* Bandwidth and sample rate requirements for wearable head impact sensors. *J. Biomech.* **49**, 2918–2924 (2016).
18. Guskiewicz, K. M. & Mihalik, J. P. Biomechanics of sport concussion: quest for the elusive injury threshold. *Exerc. Sport Sci. Rev.* **39**(1), 4–11 (2011).
19. Broglio, S. P., Lapointe, A., O'Connor, K. L. & McCrea, M. Head impact density: A model to explain the elusive concussion threshold. *J. Neurotrauma* **34**, 2675–2683 (2017).
20. Rowson, S. *et al.* Correlation of Concussion Symptom Profile with Head Impact Biomechanics: A Case for Individual-Specific Injury Tolerance. *J. Neurotrauma* **35**, 681–690 (2018).
21. Broglio, S. P., Eckner, J. T., Surma, T. & Kutcher, J. S. Post-concussion cognitive declines 472 and symptomatology are not related to concussion biomechanics in high school 473 football players. *J. Neurotrauma* **28**, 2061–2068 (2011).
22. Rowson, S. & Duma, S. M. Brain injury prediction: Assessing the combined probability of concussion using linear and rotational head acceleration. *Ann. Biomed. Eng.* **41**, 873–882 (2013).
23. Rowson, S. *et al.* Rotational head kinematics in football impacts: An injury risk function for concussion. *Ann. Biomed. Eng.* **40**, 1–13 (2012).
24. Liu, Y. *et al.* Validation and *Comparison of Instrumented Mouthguards for Measuring Head Kinematics and Assessing Brain Deformation in Football Impacts*. http://arxiv.org/abs/2008.01903. (2020).
25. Preece, S. J. *et al.* Activity identification using body-mounted sensors: A review of classification techniques. *Physiol. Meas.* **30**(4), R1–R33. https://doi.org/10.1088/0967-3334/30/4/R01 (2009).
26. Mannini, A. & Sabatini, A. M. Machine learning methods for classifying human physical activity from on-body accelerometers. *Sensors (Basel).* **10**(2), 1154–1175. https://doi.org/10.3390/s100201154 (2010).
27. Wu, W., Dasgupta, S., Ramirez, E. E., Peterson, C. & Norman, G. J. Classification accuracies of physical activities using smartphone motion sensors. *J. Med. Internet Res.* **14**(5), e130. https://doi.org/10.2196/jmir.2208 (2012).
28. Motiwale, S. *et al.* Application of neural networks for filtering non-impact transients recorded from biomechanical sensors. 3rd *IEEE EMBS Int. Conf. Biomed. Heal. Informatics* 204–207. https://doi.org/10.1109/BHI.2016.7455870 (2016).
29. Thompson, H. J., Vavilala, M. S. & Rivara, F. P. Common data elements and federal interagency traumatic brain injury research informatics system for TBI research. *Annu. Rev. Nurs. Res.* **33**(1), 1–11 (2015).
30. NIH 2020. *The Federal Interagency Traumatic Brain Injury Research Website*. https://fitbir.nih.gov/ (2020).
31. Lang, L., Moser, R. P., Odenkirchen, J. & Reeves, D. Common Data Elements. Natl. Inst. Neurol. Disord. Stroke 13: Traumatic Brain Injury (2016).
32. Broglio, S. P. *et al.* National Institute of Neurological Disorders and Stroke Department of Defense Sport-Related concussion common data elements version 1.0 recommendations. *J. Neurotrauma* **35**(23), 2776–2783 (2018).
33. Wu, L. C. *et al.* Detection of American Football head impacts using biomechanical features and support vector machine classification. *Sci. Rep.* **8**, 855 (2017).
34. Funk J., J. Crandall, M. Wonnacott, C. Withnall and D. Drive. *NFL Linear Impactor Helmet Test Protocol* (Biocore, Charlottesville, VA, Accessed Feb 1: 2019, 2017).
35. Raymond, S.J. *et al.* A deep learning approach for designed diffraction-based acoustic patterning in microchannels. *Sci. Rep.* **10**, 8745. https://doi.org/10.1038/s41598-020-65453-8 (2020).
36. Sganga, J., Eng, D., Graetzel, C. & Camarillo, D. Offsetnet: Deep learning for localization in the lung using rendered images. In *2019 International Conference on Robotics and Automation (ICRA)* 5046–5052 (IEEE (2019).
37. He, H. & Garcia, E. A. Learning from Imbalanced Data. *IEEE Trans. Knowl. Data Eng.* **21**(9), 1263–1284. https://doi.org/10.1109/TKDE.2008.239 (2009).
38. Bai, S., Zico Kolter, J. & Koltun, V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. arXiv preprint. (2018).
39. Zhang, X., Zhao, J. & LeCun, Y. Character-level convolutional networks for text classification. In *Adv. NIPS* (2015).

Acknowledgements

Author contributions

A.D., S.R., C.G., and Y.L. prepared the manuscript, S.A.Y., E.R., M.F., S.S., S.J., and I.R. performed additional data analysis I.P., N.C, O.V, A.K., B.A., A.B., T.F., W.M., S.M., N.M., P.R., H.S., L.W., and C.K., provided manuscript reviews and edits and M.Z., G.G., and D.C. provided project supervision and oversight.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to S.J.R.

Reprints and permissions information is available at www.nature.com/reprints.