Detection of magnetohydrodynamic waves by using convolutional neural networks

Fang Chen and Ravi Samtaney
Mechanical Engineering, Physical Science and Engineering Division, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia

(Dated: 16 September 2022)

Nonlinear wave interactions in magnetohydrodynamic (MHD), such as shock refraction at an inclined density interface, lead to a plethora of wave patterns with numerous wave types. Identification of different types of MHD waves is an important and challenging task in such complex wave patterns. Moreover, owing to the multiplicity of solutions and their admissibility for different systems, especially for intermediate-type MHD shock waves, the identification of MHD wave types is complicated if one relies on the Rankine-Hugoniot jump conditions. MHD wave detection is further exacerbated by nonphysical smearing of discontinuous shock waves in numerical simulations. This paper proposes two MHD wave detection methods based on convolutional neural network (CNN) to enable wave classification and identify their locations. The first method separates the output into regression (location prediction) and classification problems, assuming the number of waves for each training data is fixed. In contrast, the second method does not specify the number of waves a priori and the algorithm predicts wave locations and classifies types using only regression. We use one-dimensional (1D) input data (density, velocity, and magnetic fields) to train the two models that successfully reproduce a complex two-dimensional (2D) MHD shock refraction structure. The first fixed output model efficiently provides high precision and recall, achieving total neural network accuracy up to 99%, and the classification accuracy of some waves approaches unity. The second detection model has relatively lower performance, with more sensitivity to the setting of parameters, such as the number of grid cells $N_{\text{grid}}$ and the thresholds of confidence score and class probability, etc. The detection model achieves better than 90% accuracy with $F_1$ score $> 0.95$. The proposed two methods demonstrate very strong potential for MHD wave detection in complex wave structures and interactions.

I. INTRODUCTION

Shock waves are common features in high-speed compressible flows, where the fluid state (density, pressure, etc.) changes drastically over a thin region whose thickness is proportional to the mean free path of the gas. Because of the importance of shock waves in several applications, such as aircraft aerodynamics, astrophysical phenomena, and inertial confinement fusion (ICF), etc., the detection of shock waves is an important and challenging problem in computational fluid dynamics (CFD). Several shock detection methods have been developed for hydrodynamic shocks. Pagendarm and Seitz proposed a shock detection method based on searching the density gradient maxima. Although the method is easy to understand, proper filters must be carefully designed to remove false results. Samtaney proposed the zero-crossing of the laplacian of the density field to detect shocks. Since the normal direction of shock wave is parallel to local pressure gradient, thus normal Mach number can be obtained from pressure distribution, Lovely and Haimes proposed using the iso-surface of unity normal Mach number to represent detected shock wave surface. The above methods are relatively easily implemented, but their effectiveness and accuracy require further improvement. Kanamori and Suzuki proposed shock detection method by calculating critical lines for the vector field of the characteristics. This approach can be applied in steady and unsteady two 2D and three-dimensional 3D flows. A shock-wave-detection technique for continuum and rarefied-gas flows has been proposed using Schlieren imaging. The scheme is applicable for any existing 2D flow fields obtained experimentally or numerically. Samtaney et al. proposed to extract shocklets in turbulence simulations based on finding the minimum for a cost function defined in terms of local Rankine-Hugoniot jump conditions. Fujimoto et al. recently offered an alternative shock detection method by integrating Canny Edge Detection, which is one of the image processing methods to detect edges, and Rankine-Hugoniot relations. These two methods provide good accuracy, but implementation is somewhat complicated.

In magnetohydrodynamics (MHD), a wave is considered physical only if it satisfies both viscosity admissibility condition and evolutionary condition. In a strongly planar system, the fast and slow waves (shocks or expansion fans),

\[a)\) Corresponding mail: fang.chen@kaust.edu.sa
contact discontinuities, $1 \to 3$ and $2 \to 4$ intermediate shocks, slow and fast compound waves are considered admissible; while fast and slow waves (shocks or expansion fans), contact discontinuities and $180^\circ$ rotational discontinuity (RD) are admissible in the planar system. Note that a flow is planar if there is no derivatives in the out-of-plane (z) direction, and strongly planar if there is also a reference frame in which there is no vector component in the z-direction. Therefore, shock wave detection in MHD becomes more challenging due to the variety of wave types.

Wheatley et al. discuss in detail the variety of waves resulting from shock refraction in MHD. Most shock detection methods developed for hydrodynamics have not yet been adopted or extended to MHD due to myriad MHD wave types, which cause these methods to fail classification and detection. MHD wave classification is traditionally based on the Rankine-Hugoniot relations, producing a set of nonlinear algebraic relations and compatibility relationships for a complex wave structure. Snow et al. proposed detection method for MHD shocks based on upstream and downstream velocities relative to characteristic speeds of the system. This method was specifically applied to shock identification and classification in 2D MHD compressible turbulence-Orszag-Tang vortex evolution. Another example is that of chromospheric detections of intermediate shocks for MHD nonlinearities in sunspot atmospheres and a technique to identify the high-frequency MHD waves of an oscillating coronal loop has been developed. It is claimed that current detection methods for MHD waves detection remain somewhat limited, and thorough investigation of MHD wave detection remains largely unexplored.

This study developed two simple methods to detect MHD waves based on convolutional neural network (CNN). The proposed approach was applied to MHD shock refraction at an inclined density interface (e.g. \cite{13,14}). Fig. 1(a) shows a canonical physical set-up to investigate shock refraction. Flow is characterized by incident hydrodynamic shock sonic Mach number $M$ (or fast magnetosonic Mach number for fast mode MHD shocks); density ratio of the interface $\eta = \rho_b/\rho_0$; ratio of specific heats $\gamma$; angle between incident shock normal and interface $\alpha$; and non-dimensional strength of an initially applied magnetic field $\beta^{-1} = B^2/2p_0$, where $B$ and $p_0$ denote dimensionless magnitudes of the applied magnetic field and gas pressure. Fig. 1(b) shows that shock refraction produces a pair of reflected and a pair of transmitted waves. RF and TF are the fast reflected and transmitted magneto-sonic shocks, respectively; whereas the wave type of TS and RS is strongly dependent on the chosen parameters (e.g. $\beta$, $\alpha$ and initial magnetic field orientation whether parallel or perpendicular to the motion of incident shock, etc.) \cite{13,20}. Fig. 2 shows an example of RS and TS transitions with increasing $\beta$ (in the mesh grid is initially present) by fixing a set of parameters $(M = 2, \eta = 3, \gamma = 1.4)$. With increasing of $\beta$ in the strongly planar system (denoted as IC), RS and TS transit from a slow shock to a $2 \to 4$ intermediate shock, and then into a slow-mode compound wave $C_1$. In planar system (noted as IR), with increasing of $\beta$, RS and TS transit from a slow shock to RD followed by slow shock, and then into a RD followed by slow-mode expansion. Further details are provided elsewhere. \cite{13,20}. Although Fig. 2 is somewhat difficult to understand, it illustrates there are many different MHD shock refraction structures with different parameters that are too complicated to identify by simply extending existing methods used for hydrodynamics.

From the point of view of machine learning application to fluid mechanics, deep learning has achieved considerable...
recent success in CFD in the past few years. For the application of CNN, Lee et al. shown that developed
CNN could predict vortex dynamics at distinctive flow regimes with flow structures at different scales. Dai et al. proposed a CNN based machine learning model to predict energy conversion efficiency for multi-droplet jumping with
different initial distribution angles and radius ratios. The applicability of the machine learning based reduced order
model (ML-ROM) to three-dimensional complex flows have been examined. Specially, some shock detection methods
based on CNN have been recently proposed for hydrodynamic cases, confirming that CNN methods can produce
better detection results than some traditional methods. However, deep learning applications for MHD remain limited
to date. Kates-Harbeck et al. developed a fusion recurrent neural network (FRNN) to predict disruptive instabilities
in controlled fusion plasmas, achieving good efficiency to predict disruptions. The current study aims to develop
a simple CNN to detect MHD waves, constraining these waves to admissible types for planar and strongly planar
systems.

The outline of this paper is as follows. In Sec. II we first introduce the fixed output model in which the wave
number of each training data is fixed, followed by the detection model without fixing the number of waves for each
sample. Sec. III presents results and discussions. The two proposed models are first verified via hydrodynamic cases
to demonstrate their robustness and efficiency in Sec. III A. The application to MHD wave detection in 1D and the
reconstruction of 2D shock refraction are discussed in Sec. III B and Sec. III C, respectively. Finally, some conclusions
are presented in Sec. IV.

II. 1D CONVOLUTIONAL NEURAL NETWORKS

In this section, we present two CNN algorithms to detect the MHD waves resulting from the shock refraction
process. The two algorithms are developed based on supervised learning, using a training dataset to teach models
to yield the desired output. Hence we need to provide a labeled dataset to train algorithms to accurately classify
data and predict outcomes. The labeled dataset consists of input training datasets and labels. Primitive variables
\( U = (\rho, p, \mathbf{v}, \mathbf{b}) \) (each such state vector comprises six variables) are employed as input training dataset, where \( \rho \)
and \( p \) is the density and the gas pressure, respectively; while \( \mathbf{v} \) and \( \mathbf{b} \) are two-component velocity vector and the
magnetic induction, respectively. Each primitive variable is specified in a training sample, comprising \( N_t \) points.
Presently, we choose \( N_t = 100 \) points in each training sample. Furthermore, each training sample is obtained from a
1D numerical simulation or a straight line containing all five waves \( RF, RS, SC, TS \) and \( TF \) resulting from the shock
refraction process. Here, we only consider these five waves in MHD shock refraction since \( IS \) type and unshocked
density interface don’t vary as the parameters change, e.g. \( \beta, \alpha \) and \( \gamma \), etc. Thus, each training data has dimension
\((6 \times 100)\). The labels of each sample consist of labeled class \( c \in C \) and bounds \((s_1, s_2)\) for each wave. It leads to \( 3w \)
dimension of labels for a sample that consists of \( w \) waves. Presently, we consider \( C \) classes for seven MHD waves,
viz., contact discontinuity (CD), fast shock (FS), slow-mode expansion (Scw), slow-mode compound wave (Scw), slow shock (SS), rotational discontinuity + slow shock (RD + SS) and 2 → 4 intermediate shock (I24). The C classes comprise three waves for hydrodynamics cases: contact discontinuity (CD), shock wave (S), and expansion fans (Exp). We first test and verify the proposed algorithms in hydrodynamics (see Sec. III A), an example of training sample is shown in Fig. 3(a).

Fig. 3(b) shows 1D CNN architecture used in the MHD wave detection. The developed two CNN algorithms have the same main architecture except for the output layer, which is different between the two proposed algorithms. The details of the two CNN algorithms are presented in the subsections Sec. II A and Sec. II B. The neural network architecture is composed of the input layer, convolutional layer followed by max-pooling layer (3 levels), two fully connected layers, and the output layer. The set of some main hyper-parameters of the two algorithms are common, e.g. the optimizer “Adam” is chosen since gradient descent algorithms are basic optimizations to attain the optimized value, and the “Adam” algorithm is one of the most traditional optimizers for its robustness and effectiveness. Another hyper-parameter is the activation function “ReLU” (Rectified Linear Unit) that is set to achieve the nonlinear connection between the fully connected layers. The set of the other hyper-parameters of the two algorithms and their details are presented in the following subsections. The neural networks are implemented using the popular TensorFlow machine learning library.

A. Fixed output model

We now present the first CNN model, where the number of waves \( w \) of each training sample is fixed (\( w = 5 \) for MHD, and \( w = 3 \) for hydrodynamic shock refraction in 1D). The fixed output system models both classification and regression problems. For each wave, the system leads to a \((2 + C)\) tensor for predicting the wave location and classifying its type. The output layer is then encoded as a \( w \times (2 + C) \) tensor. Here, we briefly introduce the hyper-parameters setting for this model:

- Prediction \( w \) wave locations leads to \( 2w \) neurons at the output layer, i.e., 10 (6) neurons for MHD (hydrodynamic) wave detection. This is a regression problem and the activation function connected to the previous layer is set as the \( \text{ReLU} \) type. Output from this part is evaluated by minimizing the mean square error (mse) cost function. The metric to monitor the training process is root mean square error (rmse).

- Classification of \( w \) wave types. This is a multi-classification problem (\( C \) classes in total) resulting in \( w \times C \) neurons due to applying the \( \text{softmax} \) activation function at the final output layer. The number of neurons

![Diagram](image)
from the C-classification is 35 for MHD cases, while it is 9 for hydrodynamic cases. The traditional cost function
for multi-classification problem “categorical-crossentropy” is adopted and the “accuracy” is used as a metric.
Appendix A provides the remaining hyper-parameter settings for this method. The model can be used for some
special cases that the number of waves is known, e.g. Sod shock tube and shock refraction problems, but it is not a
true detection algorithm since the number of waves is fixed a priori rather than remaining unspecified. However, this
model can serve as a verification process due to similarity of its architecture as in the real detection model, which will
be presented in the following subsection.

B. Detection model

Recently, several mature algorithms have been developed for real-time object detection in images, including Fast
R-CNN[32], Mask-R-CNN[24], YOLO[23], etc. The proposed detection model was inspired by the YOLO algorithm.
There are two main differences between the present detection model and the fixed output model. First, we employ
divided grids over the input data, leaving the number of waves in the samples unspecified rather than fixing their
number. Hence this algorithm models detection as a regression problem, whereas the fixed output model regards
detection as two classification and regression problems. On the other hand, this algorithm models detection as a
regression problem. In the fixed output model this is regarded as two problems: classification and regression. The
main concept is as follows.

- The algorithm divides the input data into \( N_{grid} \) grids. If the center of a wave falls into a grid cell, then this
  grid cell is responsible for detecting this wave. Each grid cell predicts \( Bo \) bounding boxes that predict wave
  locations, with correlated confidence scores \( \kappa \). Presently, we set \( Bo = 1 \) as this is deemed sufficient for our 1D
  CNN algorithm.

- The confidence score \( \kappa \) reflects how confident the model is that the box contains a wave and also how accurate
  it considers the box that it predicts. We define \( \kappa = Pr(\text{wave}) \times IOU_{true}^{pred} \), where \( Pr(\text{wave}) \) is the probability
  for a wave being present, and \( IOU_{true}^{pred} \) is the intersection over union (IOU) between the predicted box and the
  ground truth. \( \kappa = 0 \) if no wave exists in that grid cell and \( \kappa = IOU \) otherwise.

- Each bounding box comprises three predictions: two predictions are the bounds for the predicted bounding
  box, denoted as \((s_1, s_2)\), and the third prediction is the confidence score \( \kappa \), representing the IOU between the
  predicted bounding box and any ground truth box.

- Each grid cell also predicts \( C \) conditional class probabilities, \( Pr(\text{Class}_i|\text{wave}) \). These probabilities are conditioned on the grid cell containing a wave. We predict one set of class probabilities per grid cell.

- We multiply the conditional class probabilities and the individual box confidence predictions,

\[
Pr(\text{Class}_i|\text{wave}) \times Pr(\text{wave}) \times IOU_{true}^{pred} = Pr(\text{Class}_i) \times IOU_{true}^{pred}.
\]

which encodes both the probability of that class appearing in the predicted box and how well the box fits the
wave. Thus, we consider that a prediction to be correct if the classification \( Pr(\text{Class}_i) \) is correct and \( IOU > 0.5 \),
for example.

This algorithm leads to the final output layer encoded as a \( N_{grid} \times (3 \times Bo + C) \) tensor. For the chosen default
set of parameters, \( Bo = 1 \) and \( N_{grid} = 10 \), the number of neurons in the output layer is 100 (60) for the MHD
(hydrodynamic) case. Since the problem is modeled as a regression algorithm, the activation function for both fully
connected and output layers is set as “ReLU” type, loss function and metric are “mse” and “accuracy”, respectively.
Appendix B details the remaining hyper-parameters for detection algorithm.

III. RESULTS AND DISCUSSIONS

A. Verification via hydrodynamics

This section presents verifications using the two proposed models for known hydrodynamic cases. In hydrodynamics,
training datasets are selected from numerical simulations of 1D shock tube[33], shock contact interaction problems[34],
etc. To generate training dataset, we set output time as 1 and randomly generate the ratio of specific heats, initial
ratio of left/right density or pressure, Mach number, etc. Only data with three waves is acceptable for the fixed model, but this is not a limitation for the detection model. Hence we select training data with only single shock and isolated contact discontinuity for the second model. Moreover, we use at least the second order precision numerical scheme or self-similar methods, with more than 200 meshes per unit length. The 5000 sample dataset is divided into training and evaluation sets at 80:20 ratio. In each sample, the number of waves is fixed at three waves, and results from left to right in the 1D solution profiles comprise a reflected shock or expansion fan \( R \), a contact discontinuity \( CD \) and a transmitted shock \( T \). Detected waves include a combination of \( S \), \( CD \), and \( Exp \). Moreover, The main metrics to analyze predicted results are accuracy, precision, recall and \( F_1 \) score,

\[
\begin{align*}
\text{accuracy} & = \frac{TP + TN}{TP + TN + FP + FN}, \\
\text{precision} & = \frac{TP}{TP + FP}, \\
\text{recall} & = \frac{TP}{TP + FN}, \\
F_1 & = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}},
\end{align*}
\]

respectively. \( TP \) and \( TN \) are true positive and negative, respectively, i.e., the number of predictions where the model correctly predicts the positive (negative) class as positive (negative); \( FP \) and \( FN \) are false positive and negative, respectively, i.e., the number of predictions where the model incorrectly predicts the negative (positive) class as positive (negative).

![Graphs showing loss function history](image)

**FIG. 4.** Metric history for the fixed model, (a): total and location loss function history, (b): individual loss function history for each distinct wave for the 1D hydrodynamic cases.
only if $\kappa > 0.5$ and the probability of class $i$ $Pr(Class_i|wave) > 0.5$.

FIG. 5. Accuracy history for the (a) fixed model and (b) detection model in 1D hydrodynamic cases.

FIG. 6. Confusion matrix of hydrodynamic cases. (a): fixed output model, (b): detection model. $S$, shock; $CD$, contact discontinuity; $Exp$, expansion fans; $Bgd$, background.

After training the models, we employ the two trained models to predict an untrained dataset. Fig. 6 shows the confusion matrices for the trained models predicting on an untrained dataset. The fixed model achieves all metrics to unity, whereas the detection model achieves mean accuracy = 0.99, precision = 1, recall = 0.99, and $F_1 = 0.995$. The detection model metrics are only slightly reduced, and hence the performance of the detection model remains excellent. Thus, we conclude that the detection model is a promising method for MHD wave detection.

B. Detection of MHD waves

This section investigates MHD wave detection for the models verified above using hydrodynamic cases. Data from MHD shock refraction in 1D are employed as the training dataset. We consider seven MHD wave types with five fixed waves in each sample for the fixed output model ($RF$, $RS$, $CD$, $TS$, and $TF$), whereas the number of waves...
in each sample remained unfixed for the detection model. We first examine the performance of the detection model with different values for \( N_{\text{grid}} \), as shown in Fig. 7(a). The evolution of average metrics is not linear as \( N_{\text{grid}} \) increases. In the studied range of \( N_{\text{grid}} \), all mean metrics exhibit a local minimum at \( N_{\text{grid}} = 20 \), with mean recall, precision and \( F_1 \) being symmetry about \( N_{\text{grid}} = 20 \). Hence the detection model achieves good performance by considering the trade-off between accuracy \( (0.9435) \) and \( F_1 \) \( (0.9696) \) at \( N_{\text{grid}} = 15 \). If a high precision \( > 0.99 \) is expected, \( N_{\text{grid}} = 10 \) is also a good option with a relative lower accuracy of 0.9112 and \( F_1 = 0.9469 \). We next investigate the influence of thresholds of class probability on the model by fixing the confidence as 0.5, see in Fig. 7(b). The model has good performance at several thresholds of class probability, and to retain the stability of the model, we finally choose the threshold of class probability of 0.5. Figure 7(c) shows there is no significant performance difference for \( \kappa \in (0.2, 0.4) \), and \( F_1 \) and accuracy reduce as confidence increases above \( \kappa > 0.4 \). Hence \( \kappa = 0.4 \) is a reasonable choice for the confidence threshold. The detection model achieve good performance for MHD wave detection by setting suitable parameters, i.e., \( N_{\text{grid}} \), and thresholds of confidence and class probability. Therefore, we set \( N_{\text{grid}} = 10 \), \( \kappa = 0.5 \) and class probability as 0.5 thresholds for subsequent detection model investigations.

Fig. 8(a) confirms that wave location prediction is the main contribution to total loss function for the fixed output model on MHD cases, consistent with hydrodynamics outcomes. The evolution of total loss function with increasing epoch is very similar to the regression component (predicting of wave locations). Fig. 8(b) shows that classification loss function approximated to \( 10^{-7} \) for the middle wave \( (CD) \) when the model is close to being well trained (epoch \( > 20 \)). This is relatively small loss compared with the other losses since \( CD \) is shocked contact for all the training datasets, and hence it is reasonable that \( CD \) loss is the smallest component. In contrast, Fig. 8(c) shows classification loss for RF and TF waves \( \approx 10^{-2} \) when the model is close to being well trained. \( CD \) is fixed type; but each RF and TF are not, and could be fast shock \( (FS) \) or slow-mode expansion fan \( (S_{\text{exp}}) \), causing slightly higher loss function.
FIG. 8. Loss function history of the fixed output model for MHD cases. (a): Total and location loss function history. (b): Loss function history of CD. (c): Loss function history of TF and RF. (d): Loss function history of TS and RS.

for each training epoch compared with CD loss function. Figure [8(d)] shows that the same reasoning applies to classification loss evolution for TS and RS waves. TS and RS could any of SS, I24, Scw, or RD depending on the initial conditions. TS and RS accuracy was not exactly unity after training due to the prevalence of multiple wave types, whereas accuracy is precisely unity for the other three waves (RF, TF, and CD), as shown in Fig. [9]. Figure [10] shows detection model overall loss tends to be stable and accuracy approximated to 0.999 after training for MHD cases. Thus, the detection model exhibits good performance for MHD cases.

Figures [11(a)] and [11(b)] show confusion matrices for predicting an untrained dataset using trained fixed output and detection models, respectively. The fixed model achieves accuracy around 0.999 with similarly high recall = 0.998, precision = 0.999, and $F_1 = 0.998$. This high $F_1$ indicates the fixed model has excellent robustness for MHD wave detection due to using shock refraction in 1D. However, the detection model fails to find several waves, particularly CD type; whereas the fixed model detects CD type waves with accuracy = 1, and smallest loss. The parameter set is the same as for hydrodynamics: $N_{\text{grid}} = 10$, confidence = 0.5 and class probability threshold = 0.5. For the detection model, the accuracy is 0.9113, the recall and precision are 0.9133 and 0.9975, respectively, and the $F_1$ score is thus 0.9469. The relative low recall may be attributed to $N_{\text{grid}}$ and confidence and class probability threshold settings. Thus, the detection model shows promise for MHD wave detection since its simple algorithm uses only 1D input data.
C. Reconstruction of the 2D wave structure

Finally, we applied the best trained detection model based on 1D input dataset to reconstruct wave structures for 2D MHD shock refraction. The 2D dataset is obtained from numerical simulations using the developed self-similar method \cite{35,36}, and subsequently split into numerous 1D datasets by taking vertical or horizontal slices from the 2D data. Each 1D dataset contains $R$, $CD$ and $T$ waves for hydrodynamics; and $RF$, $RS$, $CD$, $TS$, and $TF$ waves for MHD. Fig. 12 and refig. 13 show reconstructed 2D wave patterns for hydrodynamics and MHD, respectively. The hydrodynamic case fails to detect some waves, including $S$ and $CD$ waves in the triple point vicinity, and $CD$ waves are erroneously detected as $S$. Non-detection and incorrect classification are caused by the waves being in close physical proximity, with consequential numerical smearing due to the underlying numerical method. The same reasoning applies for non-detection and incorrect classification near reflected waves from the bottom wall. However, the simple 1D detection model still successfully reproduces the general hydrodynamic shock refraction structure. Non-detected waves and incorrect classification occur in the quintuple point vicinity for MHD cases, i.e., where all the waves intersect. In this case, $RS$ and $TS$ are $SS$ and $I24$ waves, respectively, and incorrect wave type classification occurs for these two types. Similar to the hydrodynamics case, non-detection and incorrect classification in the reflection region from the bottom wall is more severe since the wave patterns are complicated and in physical proximity with each other. All the waves are accurately detected just a small distance away from the triple or quintuple points for hydrodynamics and MHD, respectively. One way further develop a method to connect the dots to reproduce the entire
wave front for each wave. We successfully reproduce the main 2D shock refraction structure, particularly in vicinity where all the waves interact. Although the proposed model classification is not as accurate as traditional MHD wave classification methods, *i.e.*, Rankine-Hugoniot relations, it offers a reasonable alternative that is somewhat easier to implement and more straightforward to use.

![Confusion matrix of MHD case. (a): fixed output model, (b): detection model.](image)

FIG. 11. Confusion matrix of MHD case. (a): fixed output model, (b): detection model.

![Numerical density field overlaid with detected hydrodynamic shock waves resulting from the detection CNN model.](image)

FIG. 12. Numerical density field overlaid with detected hydrodynamic shock waves resulting from the detection CNN model. Red and blue points denote hydrodynamic shock and shocked contact, respectively.
IV. CONCLUSION

In summary, the present work represents the first instance detecting MHD waves based on a machine learning approach. MHD shock refraction includes many different wave types that interact with each other in a complex manner. Detecting such waves by traditional methods, such as those to detect hydrodynamic shocks, is cumbersome and error prone. The proposed concept comprises neural networks that employ simple 1D training datasets to detect multi-class MHD waves, such as intermediate shocks, compound waves, rotational discontinuities, etc. We developed separate fixed output and detection models. The fixed output model achieved accuracy, recall, and precision all up to 0.99. The main shortcoming for this model is that the wave number needs to be fixed a priori for all training datasets. Hence the model has limited applicability for general wave detection cases. However, the underlying CNN can be adapted to the detection model, which considered wave classification and position prediction as a single regression problem. The essential concept is that we divide the input dataset into \( N_{\text{grid}} \) grid cells, which enables the model to be applied to more general wave detection problems with the accompanying advantage that the number of waves need not be fixed. Each grid predicts the confidence score and class probability which are then used as conditions to decide whether the detection is correct or not. The detection model is sensitively affected by the parameters settings, such as \( N_{\text{grid}} \), and confidence score and class probability thresholds; achieving good accuracy approximated to 0.9 and \( F_1 \approx 0.95 \). However, the proposed detection model does not achieve good performance in the vicinity where many waves are in close physical proximity. This was attributed to numerical smearing of these waves, and that the model was trained with one dimensional datasets. The proposed CNN methods provide a viable alternative to the traditional Rankine-Hugoniot approach if acceptable accuracy is not extremely high.

ACKNOWLEDGEMENT

This research was supported by funding from King Abdullah University of Science and Technology (KAUST) under Grant No. BAS/1/1349-01-01.
TABLE I. Hyperparameters of the fixed output CCN model for MHD case

| Hyperparameter | Explanation | Specified or Best value |
|----------------|-------------|-------------------------|
| $\alpha_{lr}$   | Learning rate | $1.0 \times 10^{-3}$ |
| $n$             | Number of features/sample | 600 |
| $k$             | Number of output neurons | 45 |
| $n_1$           | Number of neurons at 1\textsuperscript{st} full connected layer | 128 |
| $n_2$           | Number of neurons at 2\textsuperscript{nd} full connected layer | 64 |
| $l_{conv}$      | Number of 1D convolutional layers | 3 |
| $f_{conv}$      | Number of 1D convolutional filters | 64 |
| $s_{conv}$      | Size of 1D convolutional filters | 5, 4 and 3 in sequence |
| $l_{pool}$      | Number of 1D max-pooling layers | 3 |
| $s_{pool}$      | Size of 1D max-pooling | 5 |
| Optimizer       | Stochastic gradient descent | Adam |
| $Activation_{full}$ | Activation function for full connected layers | ReLu |
| $Activation_{out}$ | Activation function for output layer | ReLu & softmax |
| $loss$          | Loss function | mse & categorical-crossentropy |
| $metric$        | Metric to control the running epoch | rmse & accuracy |

Appendix A: Hyperparameters of the fixed output CCN model for MHD case

Appendix B: Hyperparameters of the detection CCN model for MHD case

TABLE II. Hyperparameters of the detection CCN model for MHD case

| Hyperparameter | Explanation | Specified value |
|----------------|-------------|-----------------|
| $\alpha_{lr}$   | Learning rate | $1.0 \times 10^{-3}$ |
| $n$             | Number of features/sample | 600 |
| $N_{grid}$      | Number of grids | 10 |
| $Bo$            | Number of bounding boxes | 1 |
| $C$             | Number of classes | 8 |
| $k$             | Number of output neurons | \(N_{grid} \times (3Bo + C)\) |
| $n_1$           | Number of neurons at 1\textsuperscript{st} full connected layer | 256 |
| $n_2$           | Number of neurons at 2\textsuperscript{nd} full connected layer | 128 |
| $l_{conv}$      | Number of 1D convolutional layers | 3 |
| $f_{conv}$      | Number of 1D convolutional filters | 64 |
| $s_{conv}$      | Size of 1D convolutional filters | 5, 4 and 3 in sequence |
| $l_{pool}$      | Number of 1D max-pooling layers | 3 |
| $s_{pool}$      | Size of 1D max-pooling | 5 |
| Optimizer       | Stochastic gradient descent | Adam |
| $Activation_{full}$ | Activation function for full connected layers | ReLu |
| $Activation_{out}$ | Activation function for output layer | ReLu |
| $loss$          | Loss function | mse |
| $metric$        | Metric to control the running epoch | accuracy |
REFERENCES

1. Lebedev, A. Frank, and D. Ruytov, “Exploring astrophysics-relevant magnetohydrodynamics with pulsed-power laboratory facilities,” Reviews of Modern Physics 91, 025002 (2019).

2. D. Lindl, P. Amendt, R. L. Berger, S. G. Glendinning, S. H. Glenser, S. W. Haan, R. L. Kauffman, O. L. Landen, and L. J. Suter, “The physics basis for ignition using indirect-drive targets on the national ignition facility,” J. Plasma Phys. 11, 339–491 (2004).

3. Z. Wu, Y. Xu, W. Wang, and R. Hu, “Review of shock wave detection method in CFD post-processing,” Chinese Journal of Aeronautics 26, 501–513 (2013).

4. H.-G. Pagendarm and B. Seitz, “An algorithm for detection and visualization of discontinuities in scientific data fields applied to flow data with shock waves,” (1992).

5. R. Samtaney, “Visualization, extraction and quantification of discontinuities in compressible flows,” in International Conference on Computer Vision and Pattern Recognition (Hilton Head Island, SC, 2000).

6. D. Lovely and R. Haines, “Shock detection from computational fluid dynamics results,” in 14th Computational Fluid Dynamics Conference (Norfolk, VA, U.S.A., 1999).

7. M. Kanamori and K. Suzuki, “Shock wave detection based on the theory of characteristics for CFD results,” in 20th AIAA Computational Fluid Dynamics Conference (Honolulu, Hawaii, 2011) https://arc.aiaa.org/doi/pdf/10.2514/6.2011-3681.

8. M. Kanamori and K. Suzuki, “Shock wave detection in two-dimensional flow based on the theory of characteristics from CFD data,” J. Comput. Phys. 230, 3085–3092 (2011).

9. H. Akhlaghi, A. Daliri, and M. R. Soltani, “Shock-wave-detection technique for high-speed rarefied-gas flows,” AIAA Journal 55, 3747–3756 (2017) https://doi.org/10.2514/1.J055819.

10. R. Samtaney, D. I. Pullin, and B. Kosović, “Direct numerical simulation of decaying compressible turbulence and shocklet statistics,” Physics of Fluids 13, 1415–1430 (2001).

11. T. Fujimoto, T. Kawasaki, and K. Kitamura, “Simpler method of shock wave detection by using Canny method,” in 2018 Fluid Dynamics Conference (Atlanta, Georgia, 2018).

12. R. Fujimoto, T. Kawasaki, and K. Kitamura, “Canny-Edge-Detection/Rankine-Hugoniot-conditions unified shock sensor for inviscid and viscous flows,” J. Comput. Phys. 396, 264–279 (2019).

13. E. G. Falbe and S. S. Komissarov, “On the inadmissibility of non-evolutionary shocks,” J. Plasma Phys. 65, 29–58 (2001).

14. F. F. Kenzel, R. D. Blandford, and F. Coppi, “MHD intermediate shock discontinuities. Part I. Rankine-Hugoniot conditions,” J. Plasma Phys. 42, 299–319 (1989).

15. V. Wheatley, D. I. Pullin, and R. Samtaney, “Regular shock refraction at an oblique planar density interface in magnetohydrodynamics,” J. Fluid Mech. 522, 179–214 (2005).

16. B. Snow, A. Hillier, G. Murtas, and G. J. J. Botha, “Shock identification and classification in 2D magnetohydrodynamic compressible turbulence-Orszag –Tang vortex,” Experimental Results 2, E35 (2021).

17. J. Houston, D. B. Jess, R. Keppens, M. Stangalini, P. H. Keys, S. D. T. Grant, S. Jafarzadeh, L. M. McFedrice, M. Murabito, I. Ermolli, and F. Giorgi, “Magnetohydrodynamic nonlinearities in sunspot atmospheres: Chromospheric detections of intermediate shocks,” The Astrophysical Journal 892, 49 (2020).

18. P. Allian and R. Jain, “The need for new techniques to identify the high-frequency MHD waves of an oscillating coronal loop,” Astronomy & Astrophysics 650, A91 (2021).

19. R. Samtaney, “Suppression of the Richtmyer–Meshkov instability in the presence of a magnetic field,” Phys. Fluids. 15, L53–L56 (2003).

20. F. Chen, V. Wheatley, and R. Samtaney, “Magnetohydrodynamic shock refraction at an inclined density interface,” Phys. Fluids. 224, 1–27 (2022).

21. B. Colvert, M. Alsalam, and E. Kanso, “Classifying vortex wakes using neural networks,” Bioinspiration & Biomimetics 13, 025003 (2018).

22. S. L. Brunton, B. R. Noack, and P. Koumoutsakos, “Machine learning for fluid mechanics,” Annu. Rev. Fluid Mech. 52, 477–508 (2020) https://doi.org/10.1146/annurev-fluid-011619-050214.

23. T. Nakamura, K. Fukami, K. Hasegawa, Y. Nabe, and K. Fukagata, “Convolutional neural network and long short-term memory based reduced order surrogate for minimal turbulent channel flow,” Phys. Fluids. 33, 025116 (2021).

24. M. Monfort, T. Luciani, J. Komperda, B. Ziebart, F. Mashayek, and G. E. Marai, “A deep learning approach to identifying shock locations in turbulent combustion tensor fields,” in Modeling, Analysis, and Visualization of Anisotropy, edited by T. Schultz, E. Özsarlan, and I. Hotz (Springer International Publishing, Cham, 2017) pp. 375–392.

25. Y. Liu, Y. Lu, Y. Wang, D. Sun, L. Deng, F. Wang, and Y. Lei, “A CNN-based shock detection method in flow visualization,” Computers & Fluids 184, 1–9 (2019).

26. S. Lee and D. You, “Analysis of a convolutional neural network for predicting unsteady volume wave flow fields,” Phys. Fluids. 35, 035152 (2021) https://doi.org/10.1063/5.0042768.

27. L. Dai, S. Ding, S. Gao, Z. Hu, Z. Yuan, and X. Wu, “The prediction of energy conversion during the self-propelled jumping of multidroplets based on convolutional neural network methods,” Phys. Fluids. 34, 012101 (2022) https://doi.org/10.1063/5.0065360.

28. G. Novati, H. L. de Laroussilhe, and P. Koumoutsakos, “Automating turbulence modelling by multi-agent reinforcement learning,” Nature Machine Intelligence 3, 87–96 (2021).

29. S. Lee and D. You, “Analysis of a convolutional neural network for predicting unsteady volume wave flow fields,” Phys. Fluids. 35, 035152 (2021) https://doi.org/10.1063/5.0042768.

30. E. G. Falbe and S. G. Glendinning, S. H. Glenser, S. W. Haan, R. L. Kauffman, O. L. Landen, and L. J. Suter, “The physics basis for ignition using indirect-drive targets on the national ignition facility,” J. Plasma Phys. 11, 339–491 (2004).

31. R. Samtaney, “Fast R-CNN,” in Proceedings of the IEEE international conference on computer vision (2015) pp. 1440–1448.

32. K. He, G. Gkioxari, P. Dollár, and R. Girshick, “Mask R-CNN,” in Proceedings of the IEEE international conference on computer vision (2017) pp. 2961–2969.

33. I. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection,” in Proceedings of the IEEE conference on computer vision and pattern recognition (2016) pp. 779–788.

34. G. A. Sod, “A survey of several finite difference methods for systems of nonlinear hyperbolic conservation laws,” J. Comput. Phys. 27, 1–31 (1978).
35 R. Samtaney, “Computational methods for self-similar solutions of the compressible Euler equations,” J. Comput. Phys. 132, 327 – 345 (1997).

36 F. Chen and R. Samtaney, “A numerical method for self-similar solutions of ideal magnetohydrodynamics,” J. Comput. Phys. 447, 110690 (2021).