Sensor fusion with Deep Learning

Jires Donfack Voufo

***Abstract*—**

I. Motivation

Over the years the advancement in transistors technology made it possible to build computers with greater processing power.This also helped to solve more complex issues related to artificial intelligence.AI can be divided in two subsections which are machine learning and deep learning.The decision to use either of this two algorithm or to combine both in a project depend on the type of problem we are trying to solve.Different DL algorithms are getting develop and refine for safety critical applications such as pedestrian detection in vehicles.In this case the DL model has to process and fuse the data from the different sensors to make the appropriate decision related to the type of object it sees.

In this paper we will focus on deep learning with its application in the field of sensor fusion.We will define what sensor fusion is,explore the architecture of a multi-modal sensor fusion network.We will also use a practical example for illustration purpose and give advantages of DL in sensor fusion

II. Foundation

Deep learning extract patterns from data using neural networks with the aim to teach a computer how to learn a task using raw data[1]. The application of DL can be found in image recognition,speech recognition,big data and natural language processing.DL tend to perform better than other machine learning algorithm when we dealing with large amount of data.Autonomous vehicles use sensor fusion with a DL algorithm to be able to identify objects in their surroundings.Sensor fusion consist of combining data coming from multiple sensors.The aim is to get the highest possible accuracy of the parameters being measured or the environment that we are trying to monitor. To main criteria in choosing a sensor for fusion application are:how the sensors are compatible when it comes to working in the same environment and how much the data coming from those sensors can complement each other.In case we would like to have redundancy to improve the reliability of the system,we can build a fusion setup in which sensors simply duplicate information.In an application where we want to relevant fusion information from our fusion network it is important to make sure sensors complement each other when it comes to parameters such as:data rates,field of view, range and sensitivity [4]

III. sensor Fusion

A sensor is a device that detects and responds to some type of input from the physical environment.As output, it generate an analog or digital signal(voltage ,current output) that can further be process by circuitry and computer systems.sensors can measure some parameters such as temperature, pressure,vibration,acceleration. [1] A sensor is consider passive when it measures the energy coming from the environment and active when it emit it own energy and measure the reaction.Before starting with the fusion, sensors need to be calibrated and synchronized.

A. Benefits of Sensor Fusion [9]

- Reliability:
  Having more than one sensor help to cater for a potential partial sensor failure.This help to improve redundancy and reliability
- Extended spatial coverage:
  We can have each sensor covering a separate area.When we put together all the different areas covered by those sensors,we get a better coverage of surrounding environment.
- Extended temporal coverage:
  With each sensor updating at different time intervals,we can interpolate those updates put them together to improve the temporal coverage
- Increased Confidence:
  Combining sensor data will provide increased confidence by providing measurements resilient to the uncertainties in any particular sensor based on the combined coverage and error mitigation of all sensors.
- Reduced Uncertainty:
  Given the resilience of multiple sensors to the specific uncertainty of any one, the overall uncertainty of the perception system can be drastically reduced using sensor fusion.
- Robustness against noise:
  With different sensor sources,we can detect when any single sensor has come across noise.This helps to attenuate the effect of noise in the system.
- Increased Resolution:
  When we combine the data from the sensor network,we improve the resolution of measurements

The different types of sensor fusion that exist are [3]:

- Data Fusion(Data in-Data out):
  Here the fusion is done at the raw or pre-processed data level collected from sensors with the aim to improve accuracy and precision by directly combining information coming from different sources.It is also known as Early fusion.

- Feature Fusion(Data in-Feature out):
  Here we extract from each sensor the features from the raw data before combining it.This makes it possible to get a better representation of the environment.

- Decision Fusion(Feature in -Decision Out):
  After the raw data get collected and features are extracted,decisions made at each individual sensor level finally get combined.By doing so, we can improve the final decision making process.It is also known as late fusion.Decision fusion gives better performance than others because errors from multiple models are dealt with independently.

- Multi-modal sensor Fusion:
  Here we combine sense data from different physical parameters. As example we can mention the estimation of a cellphone orientation using accelerometer,gyroscope and magnetometer

- Hierarchical sensor Fusion:
  Here the fusion of information is all done at sensor,feature and decision level at the same time.

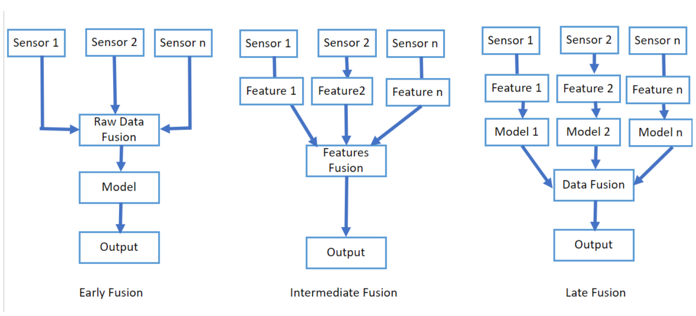

Fig. 1. Early ,Intermediate & Late fusion

Based on the way sensors in a network interact ,we can classify Sensor Fusion Algorithms as follow [4]:

- Complementary: Here the sensors work in an independent way , but it is still possible to combine their data to get a better understanding of the environment.
- Competitive: here all sensors measure the same parameter and each of them will give its own measurement.This way we can increase the reliability ,accuracy and robustness of the system.Due to redundancy in this type of setup,we have to pay attention not get inconsistent readings.
- Cooperative: In this setup independent sensors measurements are combined in a way that at the end we get an image of the environment that no single sensor alone will be able to provide. The picture bellow illustrate the working principle to different sensor network.

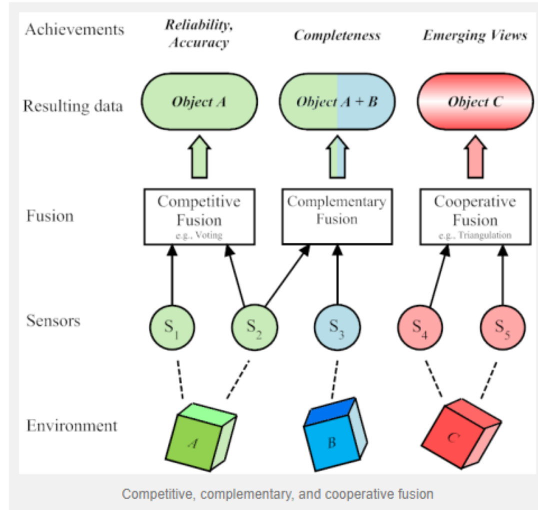

Fig. 2. Sensor Network Type [5]

IV. Deep Neural network

To be able to predict and make some decision, the deep learning models process and analyse huge amount of data during the training process.Here we will dive into the architecture and working principle of a neural network.

A. *Layers*

They represent the data processing module of the network.It can take as input one or many tensors and output also one or many tensors.Each layer has a weight and bias which are parameters that are important for the performance of a model during the training/learning process.The combination of layers form a neural network as we know.

- Input layer:first layer that is responsible for collecting the data
- Hidden Layer:positioned between the input and output layer.As example we can mention fully connected, convolutional and fully connected layer
- Output Layer: as the last layer of the network,it is there to generate the final output.

The different types of layers that can be used are:

- Fully connected layers:
  also known as dense layers, are the foundation of any feedforward neural network. They bind all the neurons from one layer to the next They are the basic building blocks of your model.
- Convolutional layers:
  These layers are essential for working with photo and video data. Using a set of convolved filters with the input image, they detect and extract the most critical features to create a detailed map of its features.
- Recurrent layers:
  Specialise for applications like text, speech, or time series. These layers maintain an updated internal state

at each time step, helping the model to remember important information from the past and make more accurate predictions.

B. Input data and Targets

The neural network maps the input data to some predictions.while the model is being trained, the possible predictions get compared to a certain ground truth with the help of the loss function.

C. Loss Function

Mathematical function to evaluate the deviation between the prediction and the actual values.During the training process of a model, the goal is to get the loss to be as low as possible.

D. Convolutional Neural network [8]

Convolutional Neural Networks (CNNs) are a type of neural network that are commonly used in image and video processing. They are designed to automatically extract and learn features from images, making them well-suited for tasks such as image classification, object detection, and image segmentation.The algorithmic capabilities of CNNs include analyzing images and videos to identify patterns and features, such as edges and shapes, and using these patterns to make predictions about the content of the image or video. CNNs can also use transfer learning, which allows them to reuse knowledge from previously trained models to improve accuracy and reduce training time.

The working steps of a CNN are as follow:

- Convolution:
  The first phase in a CNN is convolution, which involves applying a set of filters to the input image to extract features. The filters are small matrices of weights that are learned during the training process. The convolution operation involves sliding the filters across the input image and computing a dot product between the filter weights and the corresponding pixels in the input image.
- Activation:
  The second phase is activation, which involves applying a non-linear activation function to the output of the convolution operation. This is done to introduce non-linearity into the model and allow it to capture more complex patterns in the data.
- Pooling:
  The third phase is pooling, which involves downsampling the output of the convolution operation to reduce the spatial dimensionality of the feature maps. This can be done using various pooling operations, such as max pooling or average pooling.
- Flattening:
  The fourth phase is flattening, which involves flattening the pooled feature maps into a single vector. This vector can then be fed into a fully connected neural network for further processing.
- Fully Connected Layers:
  This neural network layers takes the flattened feature vector as input and generate the final output. The output layer typically uses a softmax activation function to generate a probability distribution over the possible classes or labels.
- Training:
  It is about optimizing the weights of the CNN to minimize a loss function. This is done using gradient descent and backpropagation, where the error in the output is propagated backwards through the network to update the weights. CNN works by applying convolutional filters to the input image, followed by activation, pooling, flattening, and fully connected layers. The network is trained using gradient descent and backpropagation to optimize the weights and minimize the loss function.

E. Neural Network Evaluation

Here we have to evaluate how many correct prediction the train model makes

A Deep learning model of a sensor fusion is illustrated by the picture bellow:

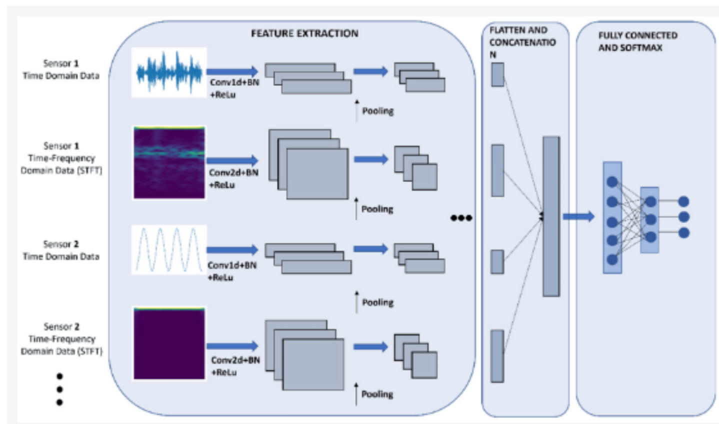

Fig. 3. Sensor Fusion model with DL [6]

V. Advantages Of using DL in Sensor fusion [7]

- classification of the relevance of each sensor to specific tasks to minimize or ignore data from sensors that are deemed to be less important.
- The fact that the DL model is continuously train helps to identify any potential changes in system behavior.
- The DL model can predict possible sources of failures, which helps with preventative maintenance.

VI. Practical case [10]

To illustrate how sensor fusion works, we use the Forward Vehicle Sensor Fusion example from Matlab.

- Steps 1:
  We started by loading the forward vehicle sensor fusion project and opening the Testbench model.Opening the model runs a script that initializes a driving scenario and automaticaly configures sensor fusion and tracking parameters.

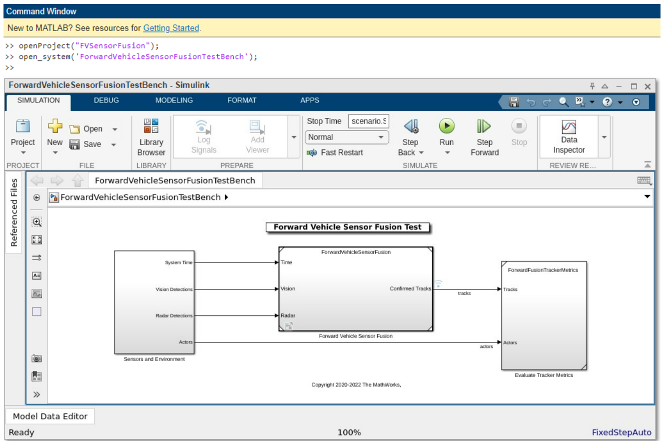

Fig. 4. step 1

- Steps 2:
  opening of the 'Sensors and Environment'.In this subsystem we can configure the road,place vehicles and synthesize sensors.

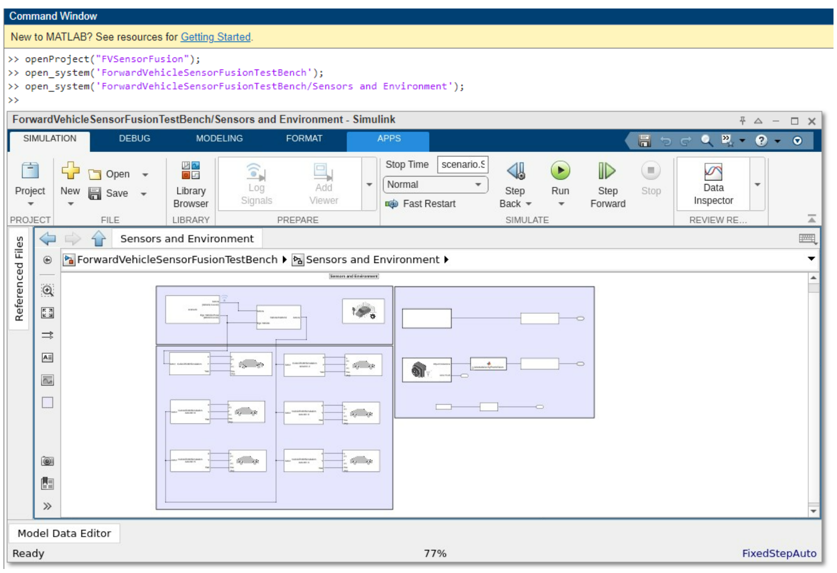

Fig. 5. Step 2

VII. Conclusion

Sensor fusion helps systems to be more consistent, accurate and dependable.For a better result the sensors need to be synchronised in time and space before sending any data to the neural network for training and predictions. Deep learning can achieve higher level of performance compare to traditional mathematical model when it comes to sensor fusion.The quality of the data also affect how well the DL model performs.

References

[1] https://www.techtarget.com/whatis/definition/sensor
[2] Bak, T. (2000). Lecture Notes: Estimation and Sensor Fusion. Department of Control Engineering.
[3] https://www.geeksforgeeks.org/concept-of-sensor-fusion-and-its-types/
[4] Bak, T. (2000). Lecture Notes: Estimation and Sensor Fusion. Department of Control Engineering.(page 5)
[5] https://netwerkt.wordpress.com/2011/03/30/the-different-types-of-sensor-fusion-complementary-competitive-and-cooperative/
[6] Kullu, O.; Cinar, E. A Deep-Learning-Based Multi-Modal Sensor Fusion Approach for Detection of Equipment Faults. Machines 2022, 10, 1105. https://doi.org/10.3390/machines10111105
[7] https://www.renesas.com/us/en/blogs/sensor-fusion-deep-learning
[8] https://www.linkedin.com/pulse/how-convolutional-neural-networks-revolutionizing-computer-pandey/
[9] https://etd.ohiolink.edu/acprod/odb_etd/ws/send_file/send accession=case1495751146601099&disposition=inline
[10] https://de.mathworks.com/help/driving/ug/forward-vehicle-sensor-fusion.html