LiDAR and Thermal Image Situational Awareness

Eyassu Mongalo, Megan Hu
Advisor: Mizanur Rahman Jewel

May 2025

Overview of what we learned

Intro - 1

- Intro
- 3Blue1Brown Neural Networks playlist
- MNIST Handwritten Digit Recognition in PyTorch
- Thermal Data
- LiDAR Point Cloud Data
- Papers read
    - Multimodal Survey
    - Where2Comm

Structure of Neural Networks

Neural Networks – 1

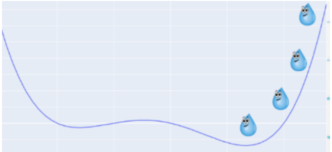
- *Multilayer perceptron Neural Networks*
  - Layers of networks
    - *Input layers*
    - *Hidden layers*
    - *Output*
- *Neurons*
  - *Activation and weights*
  - *Weighted sum of layers using activation functions*
    - Sigmoid function (aka logistic curve)
    - Inactivity bias

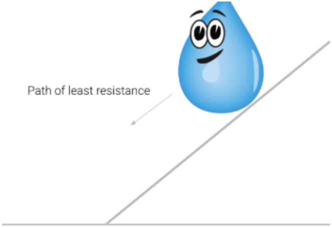
Gradient Descent

Neural Networks – 2

Motivation: How do we quantify the effectiveness of our network?

- *Neural Network Cost function*
  - Input
  - Output
  - Parameters
- *Objective: Reaching the local minimum*

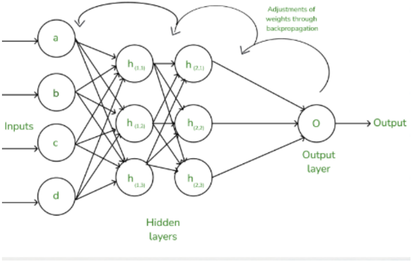




Backpropagation

Neural Networks - 3

Backpropagation is an algorithm that computes the gradient for a single training example.



However. . . Computationally expensive

Potential Solution: Mini-batch gradient descent

LLM's explained

Neural Networks – 4

- *Large language models are trained on huge quantities of data*
  - *Allows them to assign probabilities to a list of "next possible words"*
- *Pretraining*
  - *Parameters*
    - *Continuous weights that determines the probabilities assigned to words*
    - *Repeatedly refined during training*
- *RLHF (Reinforcement Learning with Human Feedback)*
  - *Workers manually flag problematic or inaccurate predictions*

Transformers

Neural Networks – 5

1. *Input is tokenized into smaller units*
2. *Tokens are converted into vectors using embeddings*
3. *Vectors are passed into the attention block*
   (focus mechanism to weigh relationships)
4. *Output is passed through a feed-forward layer (MLP)*
5. *Steps 3–4 are repeated across layers*

Attention

Neural Networks – 6

Core idea: Each token looks at other tokens to decide what matters most.

How it works:

- *Compare query to key vectors (similarity)*
- *Compute attention scores (relevance)*
- *Apply softmax* $\rightarrow$ *weighted sum of value vectors*

Softmax

Neural Networks - 7

Softmax normalizes output scores from neural networks/attention mechanism and turns them into a probability distribution

$$\text{SoftMax}(z)_i = \frac{e^{z_i}}{\sum_j e^{z_j}}$$

In attention, softmax helps a token decide how much to "pay attention" to other tokens

- Core idea: larger scores give stronger weights, smaller scores fade away

DNN, RNN, and CNN

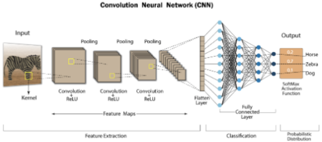
MNIST Pytorch - Prelude

- DNN - Deep Neural Network
  - The other 2 are built using this type
  - Used for more general use-cases
- RNN - Recurrent Neural Network
  - Centered More toward video and text
- CNN - Convolution Neural Network
  - Centered toward image and audio
  - This will likely be the one our project uses

Brief Overview

MNIST Pytorch - 1

- *Convolution:*
  - *Filters, Kernel, Channels*
- *MaxPool:*
  - *Splits tensor into groups of* $N \times N$, *taking the max*
  - $Size = (K//N) \times (K//N)$ *where* K *is input size*
- $ReLU(x) = max(0, x)$
- *Linear:*
  - *Node-like structure you're used to*



Backpropogation Calculus

MNIST Pytorch - 2

We define the loss function as:

$$\text{Loss} = \sum_i (A(i) - Y(i))^2$$

The derivative of the ReLU function is the unit step function:

$$\frac{d}{dt}\text{ReLU}(t) = \text{UnitStep}(t)$$

Where:

$$\text{ReLU}(t) = \max(0, t) \quad \text{and} \quad \text{UnitStep}(t) = \begin{cases} 0 & \text{if } t \leq 0 \\ 1 & \text{if } t > 0 \end{cases}$$

Install/Setup

MNIST Pytorch - 3

- What is Pytorch?
  - Python module used to built train and test neural networks
  - Flexible deep learning framework
  - Must be installed through Pip
- Torchvision
  - Companion module for computer vision tasks

Class Overview

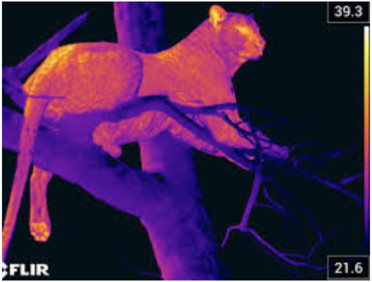
MNIST Pytorch - 4

- Net
  - Init: Sets up Convolution, Maxpool, ReLU , and Linear
  - Forward: Connects all the Convolutions and Linears
- Executive
  - Init: Downloads datasets and creates necessary items
  - Provides simple abstractions to Train, (Load/Save) Model, Test, or Print

Demo

MNIST Pytorch - 5

Thermal Data



- *Every pixel = temperature reading*
- *Applications in wildlife tracking, surveillance, autonomous vehicles, disaster response, etc.*
- *Often paired with LiDAR or RGB*

LiDAR Point Cloud Data

LiDAR/Thermal - 2

- LiDAR Point Cloud Data
  - A set of data points represented in coordinates (x, y z)
  - Used to represent 3D space
- How?
  - Points are generated using the distance and angle of laser pulses hitting a target surface

MultiModal Survey

Readings - 1

- Paper on the background of multimodal learning and how Transformers grew
  - Vanilla, Vision, and Multimodal Transformers
  - Applications and use-cases of transformers
  - Challenges and design patterns
  - Research Problems and future direction

Where2Comm

Readings – 2

Objective: Optimize multi-agent systems by improving perception

Motivating problems in multi-agent systems:

- *High communication costs*
- *Lack of spatial prioritization*
- *Fixed communication strategies*

Solutions proposed by Where2Comm framework:

- *Spatial confidence maps*
- *Confidence-aware sparse communication*
- *Multi-head attention with spatial priors and unified framework*

Plans for Next Week

Closing – 1

Focus Areas

- *Feature Extractors* – Explore new methods and read 2–3 papers on feature extraction.
- *Mini Project 2* – Complete a short AI/PyTorch tutorial to solidify new concepts.
- *LiDAR + Thermal* – Deepen understanding of feature extraction from spatial/sensor data.

Weekly Flow

Read → *Test* → *Build* → *Reflect*

Works Cited I

Closing - 2

- Sanderson, Grant. *Neural Networks*. 3Blue1Brown, YouTube. https://www.youtube.com/playlist?list=PLZHQObOWTQDNU6R1_67000Dx_ZCJB-3pi. Accessed 22 May 2025.
- Koehler, Gabriel. *PyTorch MNIST from Scratch*. Nextjournal. https://nextjournal.com/gkoehler/pytorch-mnist. Accessed 22 May 2025.
- *Thermal Imaging*. ScienceDirect Topics, Elsevier. https://www.sciencedirect.com/topics/earth-and-planetary-sciences/thermal-imaging. Accessed 23 May 2025.
- *Point Clouds for Beginners*. FARO Technologies. https://www.faro.com/en/Resource-Library/Article/Point-Clouds-for-Beginners. Accessed 23 May 2025.

Works Cited II

Closing - 2

Badrinarayanan, Vijay, et al. "Multimodal Self-Supervised Learning of Dense Representations." IEEE Xplore, 2023. https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=10123038. Accessed 25 May 2025.

Koh, Pang Wei, et al. "STRONG: Data-Efficient Image Classification via Label Propagation and One-Shot Learning." NeurIPS 2022. https://proceedings.neurips.cc/paper_files/paper/2022/file/1f5c5cd01b864d53cc5fa0a3472e152e-Paper-Conference.pdf. Accessed 25 May 2025.