Safety Helmet Detection using Computer Vision

Intro to Computer Vision: HelmNet

July 18, 2025

Contents / Agenda

Executive Summary

Project Summary

- Developed an image classification system to detect safety helmet usage using Convolutional Neural Networks (CNNs) and transfer learning.
- Explored multiple deep learning models, including a custom CNN and several versions of VGG-16 with progressively advanced architectures.
- Evaluated models on performance metrics like accuracy, recall, and F1-score to prioritize safety-critical detection.

Executive Summary

Actionable Insights & Recommendations

1. **Deploy the final model in high-risk environments** (e.g., construction sites, factories) using a camera feed + real-time inference pipeline.
2. **Expand the detection scope** to include other personal protective equipment such as vests, gloves, and eyewear using multi-class classification.
3. **Continue augmenting the training dataset** with new images captured from actual deployments to further improve model robustness.
4. **Monitor and retrain periodically** to adapt to new environments and edge cases, ensuring continued effectiveness.
5. **Integrate with alert systems** to notify supervisors instantly if a non-compliant worker is detected.

Business Problem Overview and Solution Approach

Problem:

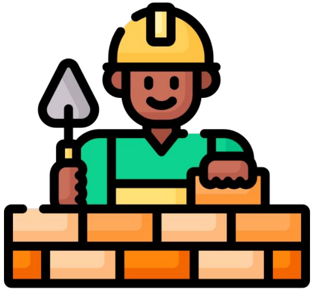
- Ensuring workplace safety in hazardous environments (construction, industrial sites).
- Manual monitoring of helmet usage is error-prone and unscalable.

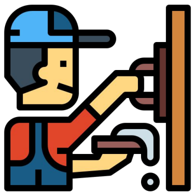
Solution Approach:

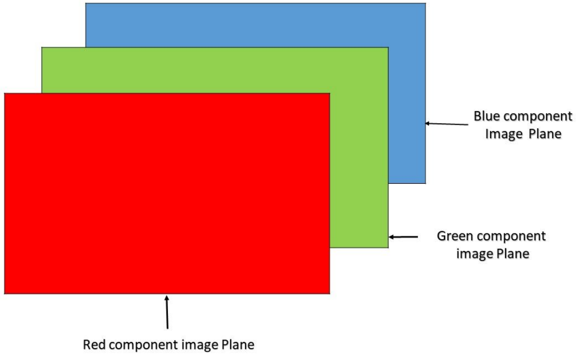
- Build a binary image classifier using deep learning to detect whether workers are wearing helmets.
- Use CNNs and VGG-16 with augmentation to improve accuracy and generalization.
- Select best-performing model for potential deployment in real-time camera systems.

Data Overview

- **Dataset:** 631 labeled images
  - 311 with helmets
  - 320 without helmets
- **Environments:**
  - Factories
  - construction sites
  - various lighting/postures
- **Input Image Size:** 200×200×3 (RGB)
- **Shape of images:** (631, 200, 200, 3)
- **Shape of labels:** (631, 1)







[RGB image is a three layered image]

Exploratory Data Analysis

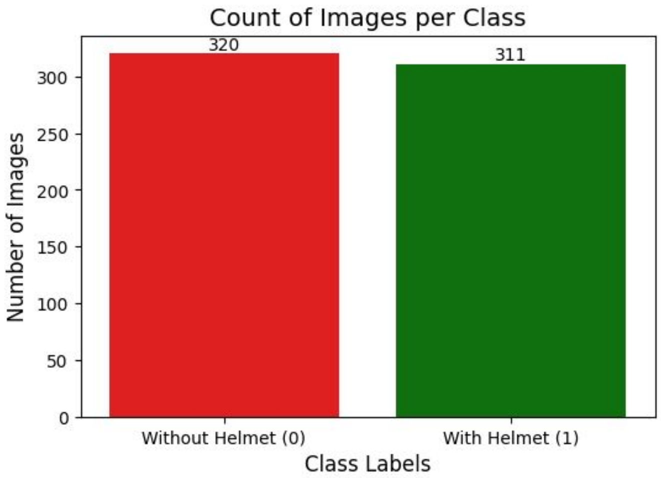
- **Plotted 2 random images:** 1 with helmet, 1 without

Worker WITH Helmet

Worker WITHOUT Helmet

Exploratory Data Analysis

Class balance: Nearly equal (311 vs 320)



Key Observations:

- Variation in lighting, pose, camera angles
- Helmet color/position may vary
- Workers engaged in different activities (standing, using tools, etc.)

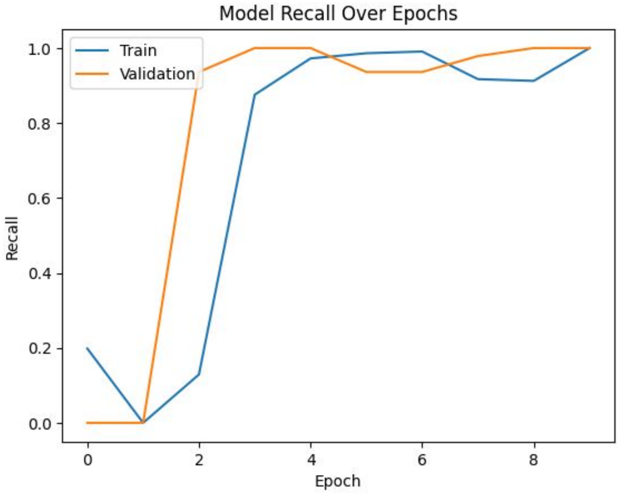
Data Pre-processing

- Converted RGB → Grayscale using OpenCV
- Plotted before and after grayscale conversion
- Split into train (70%), validation (15%), test (15%) using stratification
- Applied normalization (divided pixel values by 255.0)

Basic Convolutional Neural Network

Model Configuration

- Layers: 3 Conv + MaxPooling → Flatten → Dense → Output
- Optimizer: Adam, LR = 0.001
- Loss: Binary Crossentropy
- Epochs: 10
- Batch Size: 32

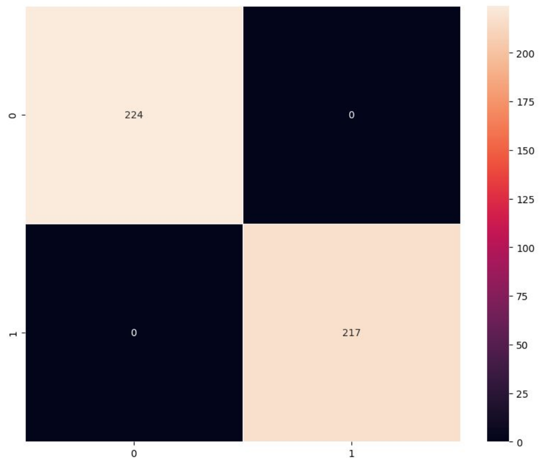


Basic Convolutional Neural Network

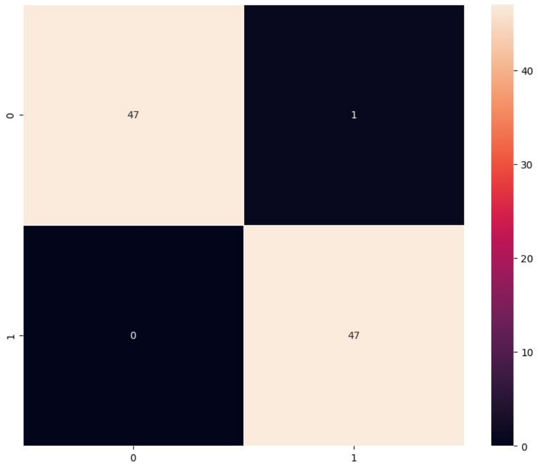
Performance

- 1 misclassification (false positive)
- Strong baseline, no overfitting

Train Recall: 100%



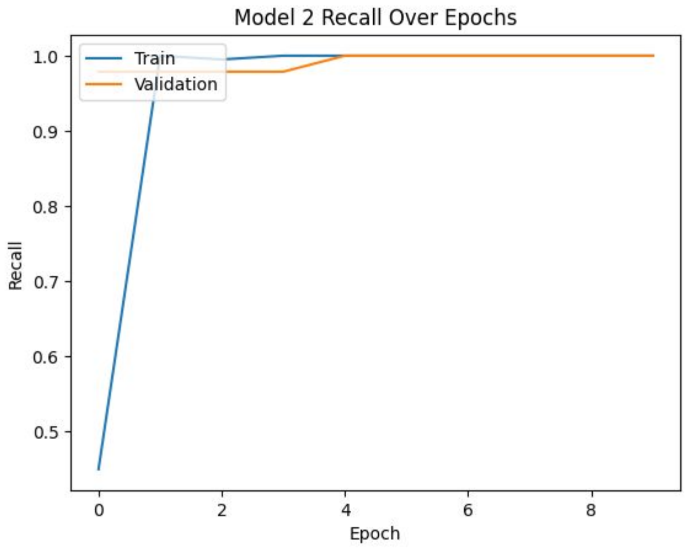
Val Recall: 98.9%



VGG-16

Model Configuration

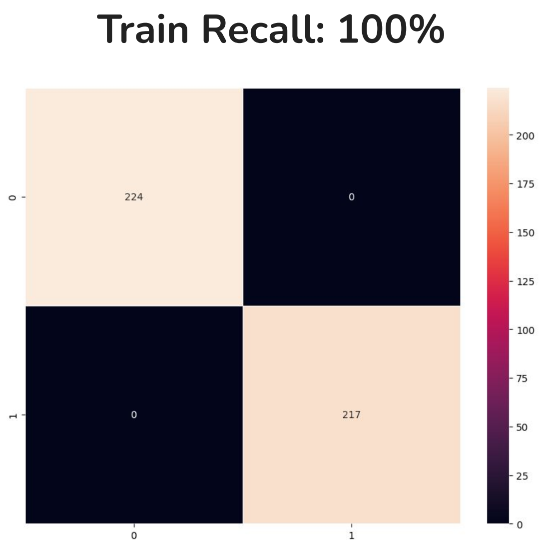
- VGG-16 base + Flatten + Dense(1, sigmoid)
- Pretrained weights (frozen)
- Optimizer: Adam, LR = 0.0001
- Epochs: 10, Batch Size: 32

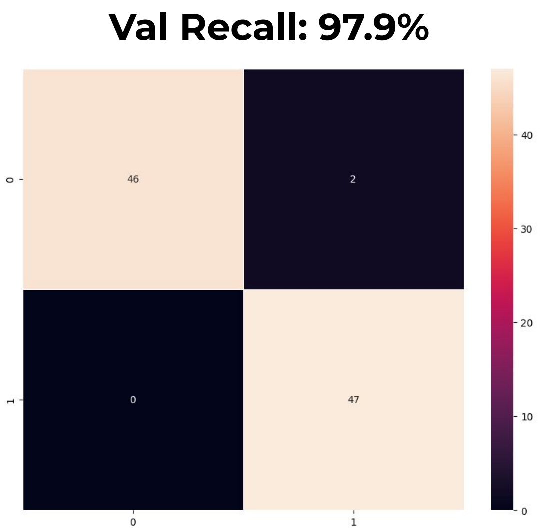


VGG-16

Performance

- 2 misclassifications
- Slightly underperforms compared to Basic CNN Model

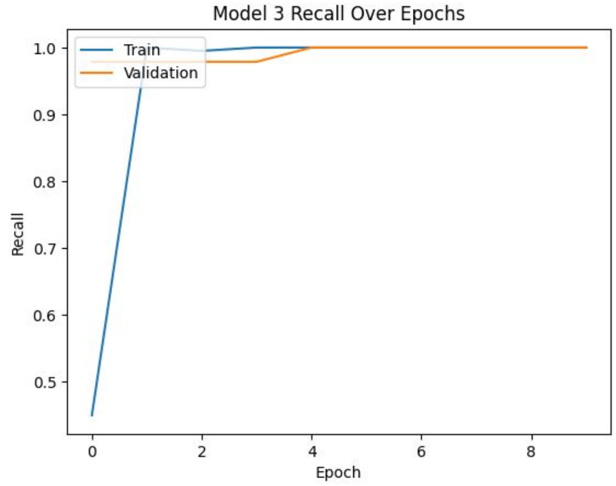




VGG-16 with FFNN

Model Configuration

- VGG-16 base + Flatten + Dense(256 → Dropout → 64 → 1)
- Optimizer: Adam, LR = 0.0001
- Epochs: 10

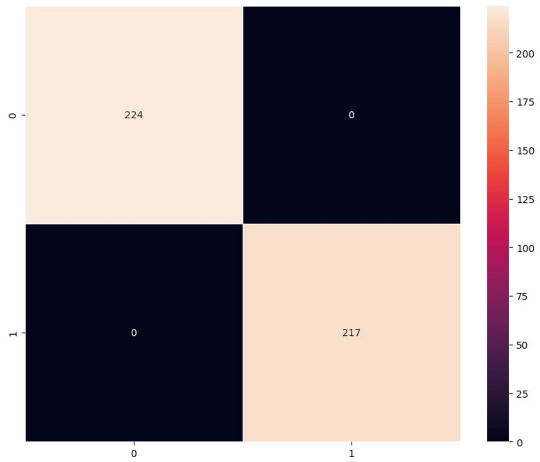


VGG-16 with FFNN

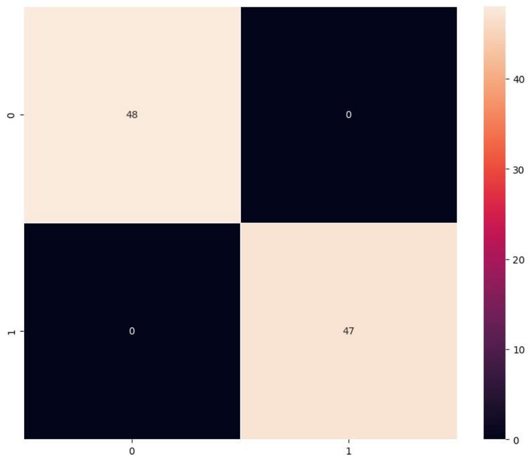
Performance

- Perfect Classification
- No overfitting
- More robust than VGG-16 Model

Train Recall: 100%



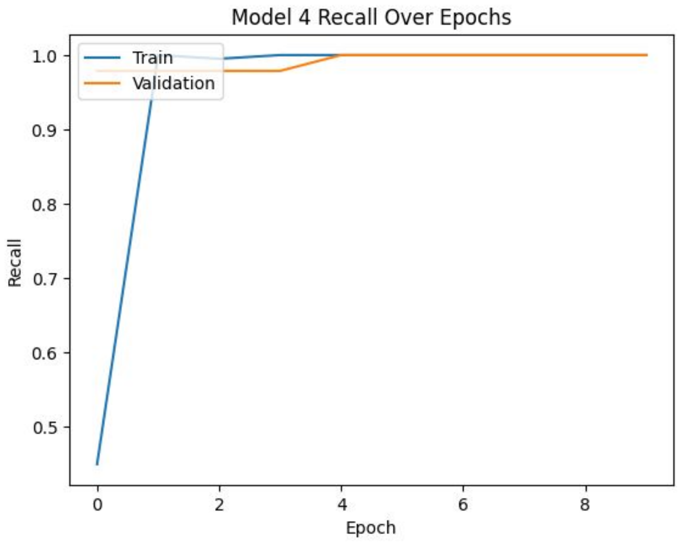
Val Recall: 100%



VGG-16 with FNN & Data Augmentation

Model Configuration

- Same as VGG-16 with FNN Model + ImageDataGenerator:
    - Rotation: 20°
    - Width/Height shift: 0.1
    - Zoom: 0.1
    - Horizontal Flip: True

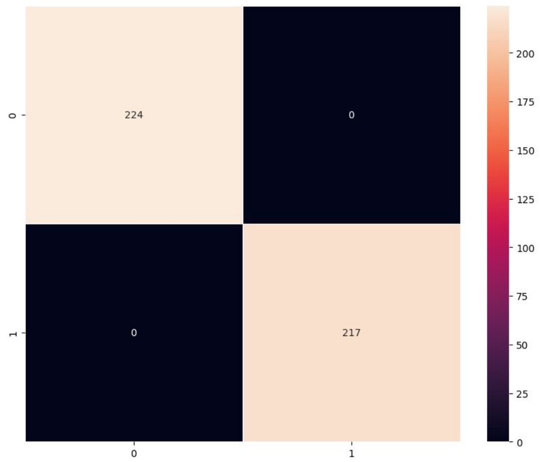


VGG-16 with FNN & Data Augmentation

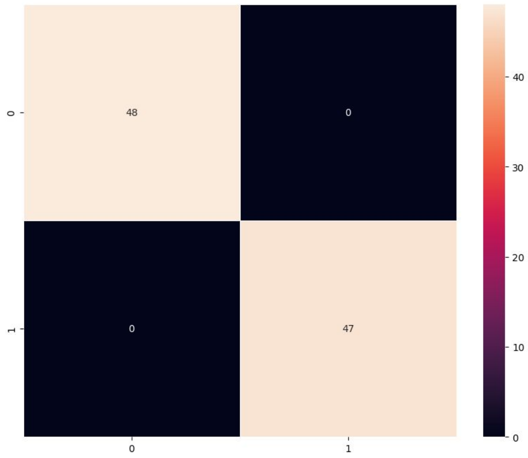
Performance

- Perfect Classification
- No overfitting
- Improved Generalization & Robust to Noise & Variation

Train Recall: 100%



Val Recall: 100%



Model Performance Comparison and Final Model Selection

Model Name	Val Accuracy	Val Recall	Misclassifications
Basic CNN	98.9%	98.9%	1
VGG-16 (Base Model)	97.9%	97.9%	2
VGG-16 + Feedforward Neural Network (FFNN)	100%	100%	0
★ VGG-16 + FFNN with Data Augmentation	100%	100%	0

Recommendation:

- VGG-16 with FNN & Data Augmentation combines the power of transfer learning, deep classification layers, and the real-world resilience provided by augmentation — making it the clear final selection for both accuracy, generalizability, and real-world deployment.

Model Performance Comparison and Final Model Selection

Superior Generalization:

- While VGG-16 with FNN also reached perfect validation performance, VGG-16 with FNN & Data Augmentation uniquely benefits from data augmentation, making it more robust to real-world image variation such as lighting, pose, angle, and background noise.

Handled Known Edge Cases:

- VGG-16 with FNN & Data Augmentation successfully corrected errors made by previous models, showing improved sensitivity to subtle visual differences in helmet usage.

Reduced Overfitting Risk:

- Data augmentation enriched the training set diversity without collecting new data, helping the model avoid overfitting and ensuring strong performance on unseen examples.

Production-Ready and Scalable:

- The VGG-16 with FNN & Data Augmentation Models's robustness and consistency make it a reliable choice for deployment in safety-critical systems, where false negatives (failing to detect someone without a helmet) can have serious consequences.

Great Learning

Power Ahead!