Project Title

Bloodstain Classification in Forensic Analysis Using Optimized 3D CNN

Prepared by

Md Al Amin Tokder

Outline

- Project Summary
- Problem Statements
- Objectives
- Project Description
- Methodology
- Result Analysis
- Implication on the economy / environment
- Ethical consideration
- Feasibility and Sustainability Analysis
- Budget Justification

Project Summary

Blood stain detection is essential in crime scene analysis as it provides valuable insights into the events that transpired, aids in identifying individuals involved and supports the investigation process of criminal cases. It helps investigators understand what happened, who was involved and when it occurred. Advancements in forensic science, particularly bloodstain analysis, have become imperative for enhancing crime scene reconstruction. Conventional methods like DNA analysis, chemical analysis often takes more time to identify bloodstain. Instead of DNA analysis this research contributes to the advancement of forensic science by introducing an innovative approach to bloodstain identification and classification by using a 3D CNN model utilizing the capabilities of Hyperspectral Imaging. Here we introduce an optimized 3D CNN with mish activation function and finding the best accuracy. This work will help to make faster investigations of forensic scene analysis and analyzing criminal cases.By applying the optimized 3D CNN model we get 97% accuracy which is higher than the existing 3D CNN (95%) and hybrid CNN(96%) model and parameters are reduced from 291895 to 125943.

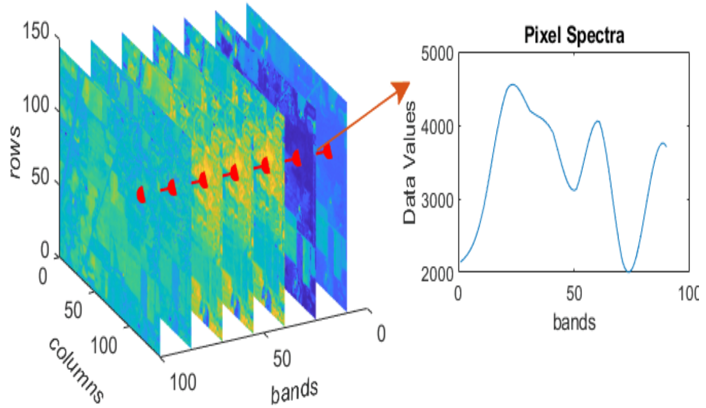

Figure-1 : Hyperspectral Image

Problem Statement

- Need Efficient feature extraction technique
- Need a custom model for better accuracy
- Forensic investigations heavily rely on the identification and classification of bloodstains
- Traditional methods for bloodstain analysis in forensic science face limitations in terms of accuracy and efficiency.

Objective

- To create a efficient method to identify substance classification in forensic science .
- To get better accuracy
- To capture both spectral and spatial information for analysis purpose
- Dimensionality Reduction

Project Description

- Hyperspectral image data is like a three-dimensional cube where each point represents a pixel in an image.It contains information about both the color and location of each pixel.
- This cube is made up of width, height and depth, where depth represents the number of different colors or spectral bands captured in the image. Each pixel is labeled with a vector indicating its category, with similarities and differences existing
- Within each category. The image may contain overlapping categories, making it challenging to distinguish them accurately.
- To simplify this complex data, preprocessing is done using Principal Component Analysis (PCA) , reduced the number of colors while preserving important spatial information, thereby aiding in analysis and interpretation.
- To effectively process hyperspectral imaging (HSI) data for image classification, a three-dimensional Convolutional Neural Network (3D CNN) is employed, utilizing its unique structure tailored for analyzing spatial-spectral information.

- First of all, the HSI cube is divided into small overlapping 3D patches to analyze individual pixels closely. These patches enabling localized analysis centered around individual pixels.

- These patches are then fed into the 3D CNN, which comprises multiple layers of 3D convolutional operations. Unlike traditional 2D CNNs, which analyze spatial information only, the 3D CNN operates on both spatial and spectral dimensions simultaneously, enabling it to capture complex spatial-spectral patterns inherent in HSI data. The structure of the 3D CNN includes multiple layers of 3D convolutional kernels followed by activation functions to introduce nonlinearity.

- This data set is available online (www.kaggle.com) and experimented on an online platform Kaggle.com . Kaggle provides free access to NVIDIA TESLA P100 GPUs. Kaggle offers a generous allocation of cold storage, providing users with 358.27 GB of storage space. Additionally, Kaggle generously allocates a substantial amount of Random Access Memory (RAM), providing users with 25 GB of memory to efficiently handle large datasets and complex computational tasks.

- This experiment is divided into three separate sets (Train/Validation/Test).These sets are Training (1690, 9, 9, 15,1) Validation (1690, 9, 9, 15, 1) Test (30424, 9, 9, 15, 1) .
- The whole dataset is spilt into three ratio . Here Training : Validation : Test is 5% : 5% : 90% .For training purpose 5% is used and 5% is used for validation purpose and rest of the 90% is considered as test set .

Proposed Methodology

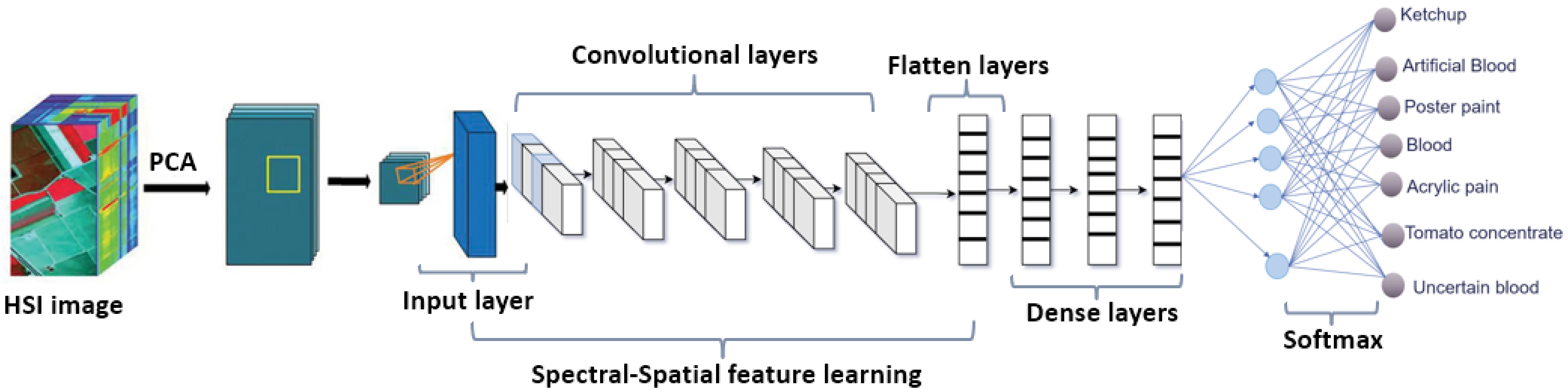

Figure-2: Optimized 3D CNN

Optimized 3D CNN [1]

Number of Layers :

Convolution Layer: 4
Flatten layer:1
Dense layer:2
Output layer 1

Activation Function uses in layers:

Convolution Layer: Mish
Dense layer: Mish
Output layer : Softmax

Layer (type)	Output Shape	Param #
input_4 (InputLayer)	[(None, 9, 9, 15, 1)]	0
conv3d_12 (Conv3D)	(None, 7, 7, 9, 8)	512
conv3d_13 (Conv3D)	(None, 5, 5, 5, 16)	5776
conv3d_14 (Conv3D)	(None, 3, 3, 3, 32)	13856
conv3d_15 (Conv3D)	(None, 1, 1, 1, 64)	55360
flatten_3 (Flatten)	(None, 64)	0
dense_9 (Dense)	(None, 256)	16640
dense_10 (Dense)	(None, 128)	32896
dense_11 (Dense)	(None, 7)	903

Total params: 125943 (491.96 KB)
Trainable params: 125943 (491.96 KB)
Non-trainable params: 0 (0.00 Byte)

Figure-3: Modified 3D CNN Model [1]

Classifiction Result Analysis

	precision	recall	f1-score	support
blood	0.99	0.97	0.98	2668
ketchup	0.99	0.99	0.99	5259
artificial blood	0.93	0.97	0.95	5774
poster paint	1.00	1.00	1.00	5963
tomato concentrate	0.95	0.92	0.94	3541
acrylic paint	0.99	0.99	0.99	7219
accuracy			0.98	30424
macro avg	0.98	0.97	0.98	30424
weighted avg	0.98	0.98	0.98	30424

Classifier Performance :

Classifiers	Accuracy	Macro avg	Weighted avg	**Parameter**
3D CNN	0.95	0.95	0.95	291895
Optimized 3D CNN	0.97	0.96	0.97	125943

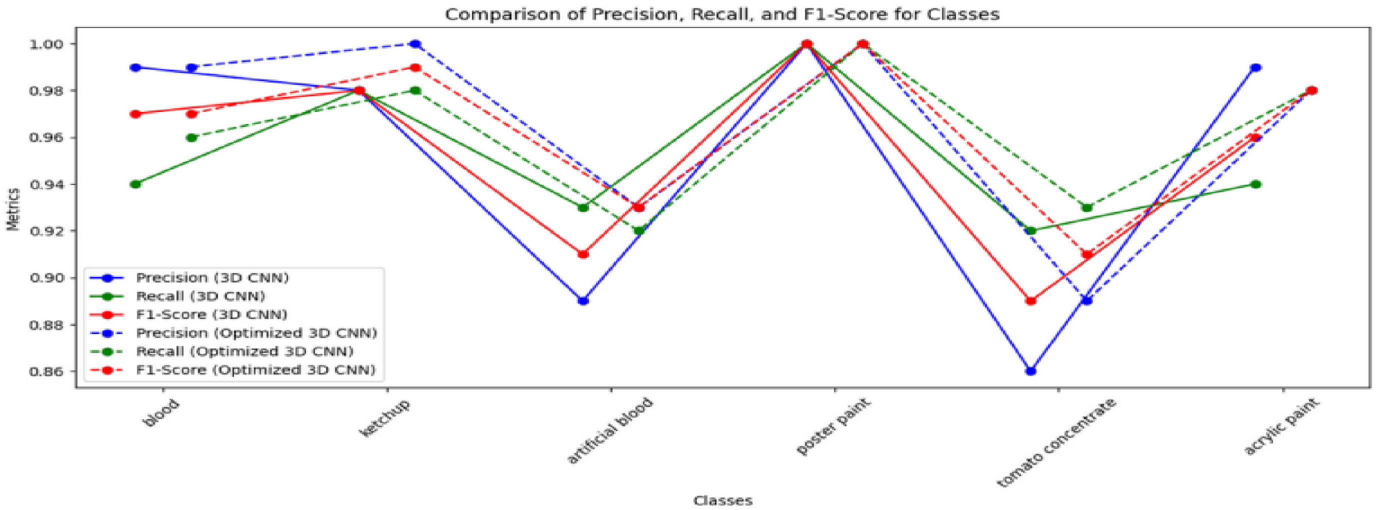

Fig. 7: Class wise Precision,Recall and F1-score Curve

Train Vs Epoch curve for modified 3D CNN :

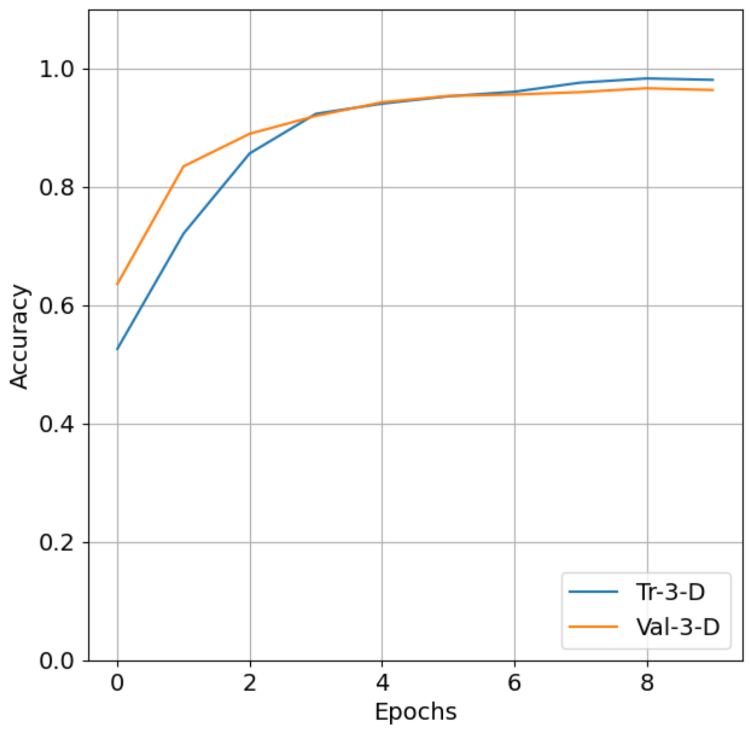

Figure-4: Accuracy vs Epochs curve

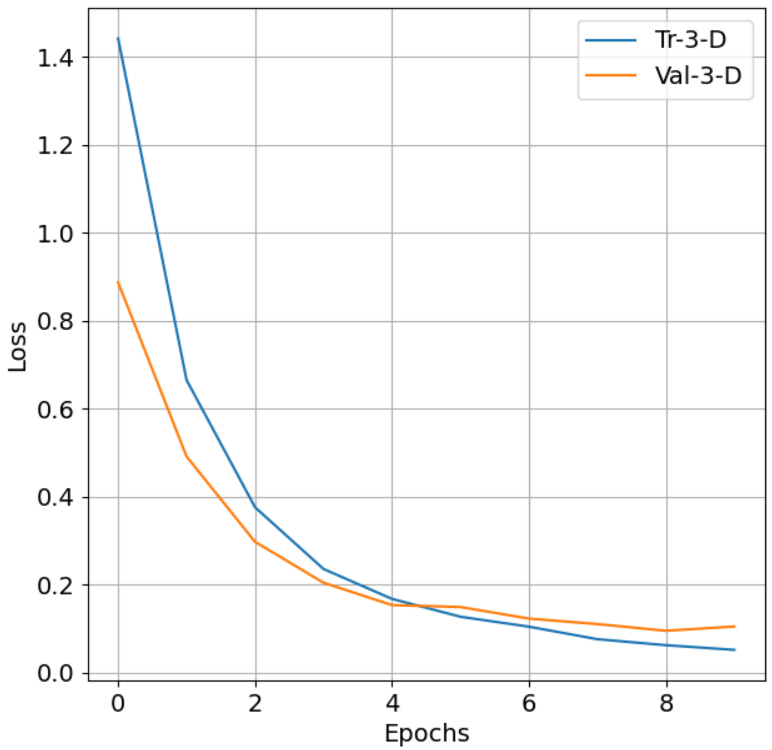

Figure-5: Loss vs Epochs curve

Accuracy Curve

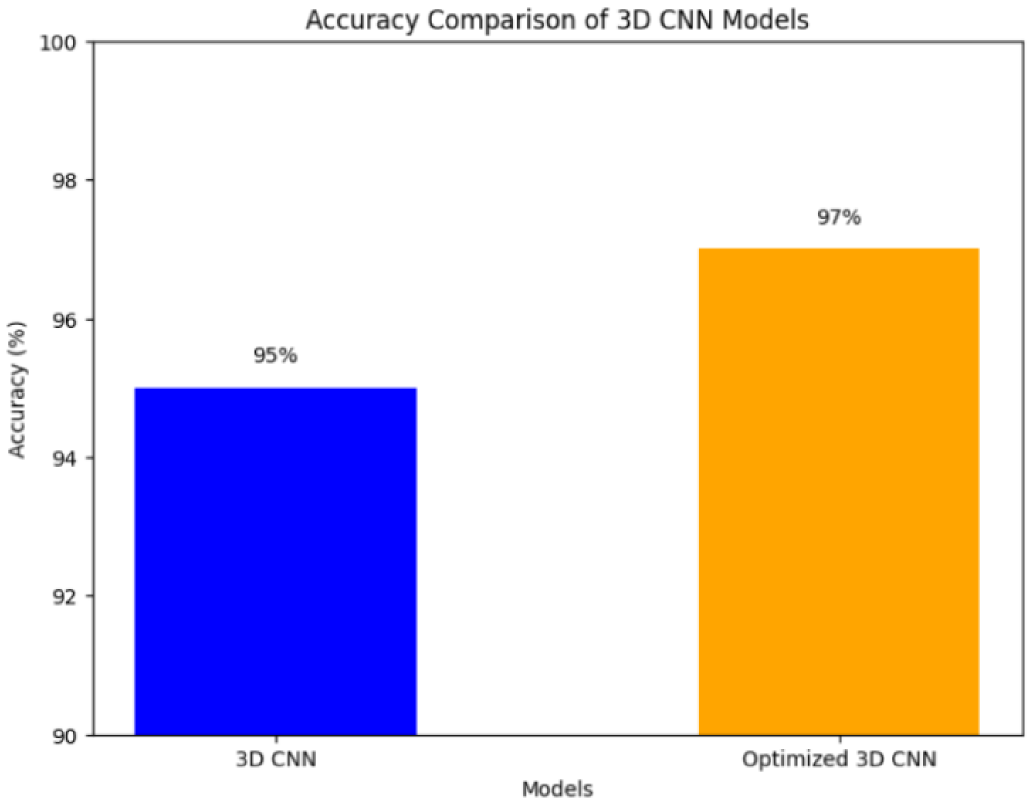

Fig. 8: Overall Accuracy Comparison

Implications on the Economy and Environment

Economy:

•**Cost-effectiveness:** Using a 3D CNN model for blood stain detection can potentially be more cost-effective than traditional DNA and chemical analysis methods. The reduction in analysis time could lead to lower labor costs and expedited legal processes.

•**Resource Allocation:** By making forensic investigations faster, resources can be reallocated to other pressing needs within the justice system or forensic departments.

Environment:

•**Reduced Chemical Use:** If the new method reduces reliance on chemical analyses, it could lessen the environmental impact associated with disposing of chemical reagents.

Ethical Considerations

- **Privacy and Data Security:** The use of advanced imaging and AI technologies raises concerns about the storage and handling of potentially sensitive biometric data.

- **Accuracy and Reliability:** While the model boasts a 97% accuracy rate, ensuring it maintains high accuracy and reliability in diverse real-world situations is crucial to prevent wrongful implications or missed detections in criminal cases.

Feasibility and Sustainability Analysis

Feasibility:

•**Technical Feasibility:** The application of a 3D CNN model utilizing Hyperspectral Imaging is technically advanced but requires specific expertise and infrastructure, which could limit its adoption to well-funded forensic departments.

•**Adoption Barriers:** There may be resistance from forensic professionals due to the shift in methodology and the need for training in new technologies.

Sustainability:

- **Ongoing Development:** Sustaining the effectiveness of the technology requires continuous updates and adaptations to new forensic challenges and evolving crime scene scenarios.
- **Scalability:** The model's adaptability to different scales of forensic operations and various types of crime scenes needs consideration to ensure it can be widely used.

Budget Justification

- **Cost Savings:** As we say earlier that DNA analysis is need more money and more costly and more time needed . Here we introduce a more efficient model that likely consumes less computational resources, leading to potential cost savings in hardware and energy consumption.

- **Investment Justification:** The improved accuracy and speed of investigations provide strong justification for investment in this technology. By expediting case resolutions, it can bring about cost savings in the judicial system and reduce the emotional and societal costs of prolonged criminal investigations.

Thank You