

Particle Classification of Electromagnetic Clusters Using the sPHENIX Detector

A Thesis Presented for
The Bachelor of Science in Physics
Degree
The University of Tennessee, Knoxville

Fredrick J Melhorn

May 2024

© by Fredrick J Melhorn, 2024
All Rights Reserved.

This project is dedicated to my family, whose support, encouragement, and appreciation for learning has been ever-present throughout my academic career.

Acknowledgements

I would like to thank my research advisor, Dr.Christine Nattrass, for initiating me into the project, and offering her guidance and advice throughout the research process. The work that I have done under her has allowed me to establish myself in physics research, and I will always be grateful to her for providing me with an enjoyable, enriching experience. I would also like to extend my gratitude for the work that she did not just with myself, but with many other undergraduates completing theses this semester.

I also want to express my gratitude to the members of my research group, including Tanner Mengel, Patrick Steffanic, and Dr. Sookhyun Lee for their assistance in the research process. Their guidance in my work has been indispensable in allowing me to develop technical skills as a researcher. The effort and care that they have put into my research experience has been remarkable. Collaborating with them has been a great enjoyment.

Abstract

This project aims to investigate the performance of a dual-branch CNN/DNN machine learning model to classify photon signatures collected by the Electromagnetic Calorimeter at sPHENIX. Neutral pions are commonly produced particles in experiments conducted at the Relativistic Heavy Ion Collider (RHIC). Upon their decay, photons are produced at scattering angles which are inversely related to the momentum of the parent neutral pion. Photons resulting from high-momentum pion decay create ill-defined clusters on the EM calorimeter. The machine learning model is tasked with sorting between merged signatures and isolated signatures. It is evident that the machine learning model is superior to the traditional sorting method, having accuracy improvements of up to 12% and 20% for merged and isolated signatures, respectively.

Contents

List of Tables	viii
List of Figures	ix
1 Introduction	1
1.1 The Quark Gluon Plasma	1
1.2 π^0 Measurements	2
2 RHIC and the sPHENIX Detector	4
2.1 Relativistic Heavy Ion Collider	4
2.2 sPHENIX Detector	5
2.3 Electromagnetic Calorimeter	7
3 Analysis Method and Data	10
3.1 Analysis Goal	10
3.2 Convolutional Neural Networks	10
3.3 Model Architecture	11
4 Model Performance and Comparisons	16
4.1 ML Model Performance	16
4.2 Comparisons of Performance	20
5 Discussion	23
5.1 Energy and Whitespace Correlations	23

5.2 Physical Implications	26
Bibliography	27

List of Tables

List of Figures

2.1	Aerial layout of RHIC	5
2.2	sPHENIX Detector and Components (1).	6
2.3	Cross Section of sPHENIX Calorimeters (1).	8
2.4	EM Calorimeter Blocks in sPHENIX (1).	9
3.1	Model Architecture	12
3.2	Merged and Isolated Sample Examples	13
4.1	Algorithm Training Process Characteristics	17
4.2	Incorrectly Predicted Merged Samples	18
4.3	Incorrectly Predicted Isolated Samples	19
4.4	Correctly Predicted Samples	20
4.5	Confusion Matrix for Traditional Method (Left) compared to the confusion Matrix for ML Model (Right).	21
4.6	Receiver Operating Characteristic (ROC) Curve	22
5.1	Accuracy of ML Model as a Function of Sample Energy	24
5.2	Average Whitespace vs Energy for Merged Samples (Left) and Average Whitespace vs Energy for Isolated Samples (Right).	25

Chapter 1

Introduction

1.1 The Quark Gluon Plasma

The Quark Gluon Plasma is a theorized form of nuclear matter that has energy densities of such high magnitudes that quarks and gluons, collectively known as partons, become free from their usual confinement in hadrons. (2). Analysis of the creation of QGP has shown that the transition of hadronic matter into the plasma occurs at a critical temperature $T_c \approx 170$ MeV (3; 4). For temperatures around T_c , the energy density of the matter rises exponentially, reaching $1\text{GeV}/\text{fm}^3$ (5)

The quarks and gluons remain conjoined, despite their asymptotic freedom from the hadrons that would normally encapsulate them. This arrangement gives the QGP characteristics that identify it as a near-perfect fluid. Upon formation, the QGP immediately begins to expand and flow, behaving as a viscous hydrodynamic fluid (3) with the lowest shear viscosity to entropy density, $\frac{\eta}{s}$, ever observed.

The QGP is a subject of interest for many different fields of physics. In the field of nuclear physics, the QGP presents an opportunity to establish a phase diagram of Quantum Chromodynamic (QCD) matter in terms of its temperature and baryon doping. Baryon doping refers to the state of having a higher ratio of quarks to antiquarks, and is often represented by μ_B . One question that is currently being

investigated is if there exists a critical value of μ_B at which phase changes occur between liquid QGP and hadronic matter. To construct a phase diagram in terms of μ_B and temperature would be a great improvement to working theories of QCD matter (2). In addition to being a subject of interest in nuclear physics, the QGP is also studied in the astronomical sciences, particularly cosmology. It is believed that QGP was the most common form of matter in the universe immediately following the Big Bang, before enough expansion and cooling had taken place for quarks and gluons to organize into more complex hadronic matter. Achieving a greater understanding of the QGP contributes directly to the understanding of the initial conditions of the early universe (2).

1.2 π^0 Measurements

Neutral pions, π^0 , are charge-neutral mesons consisting of a quark and an accompanying antiquark. Neutral pions have a very short mean lifetime of $8.5e-17$ seconds, decaying with the electromagnetic force to produce two photons. These photons scatter from the decay point at angles inverse to the momentum of the parent neutral pion at decay. These photons will propagate through space until they encounter the electromagnetic calorimeter, where they will leave an energy signature. As of now, one issue that persists in the analysis of these energy signatures has to do with those left by the photons created by high momentum neutral pions. As mentioned, the angle between these photons is very small, causing a high degree of overlap with their energy signatures on the detector. These “merged” signatures are not analytically viable with the current analysis methods employed at sPHENIX.

As of now, data analysis procedures for the sPHENIX EM calorimeter include the shower shape method. This method incorporates three stages to the data analysis procedure:

- Cluster Search
- Local Maxima Identification
- χ^2 Test

The first operation to be completed by the algorithm is a cluster search. A cluster is defined to be a distribution of energy signatures above some value spread across a group of calorimeter towers. The second operation in the algorithm is a search for the local maxima of the shower distribution upon the calorimeter towers. Local maxima are physically significant because they likely indicate a central position of the electromagnetic shower. After clusters and local maxima are found, the χ^2 test is performed to resolve the signatures recorded by the detector towers (6). Requiring $\chi^2/ND < 3$ successfully identifies 1 GeV photons with 99 % efficiency. This corresponds to applying a rather loose p-value cut of 0.01 for isolated photons.

Should particle signatures overlap, precise energy measurements and impact coordinates for each particle are lost. These characteristics can therefore only be estimated from merged signatures.

In the case of neutral pion decay, phase 2 of the algorithm outlined above becomes ineffective when the momentum of the pions approaches approximately 12 GeV/c. Once this momentum value is reached, nearly all samples recorded by the detector are composed of merged signatures, making the data available from the detector less useful due to the loss of accuracy associated with this type of signature.

Chapter 2

RHIC and the sPHENIX Detector

2.1 Relativistic Heavy Ion Collider

The Relativistic Heavy Ion Collider [2.1](#) (RHIC) is a particle accelerator located at Brookhaven National Lab which is used to collide both protons and heavy ionized nuclei at speeds near the speed of light [\(7\)](#). [\(7; 8; 9; 10\)](#). Heavy ion collisions of this scale are used to study the partonic interactions between the quarks and gluons, which make up the colliding nuclei. The temperatures reached during $Au + Au$ collisions at RHIC can exceed 300 MeV with energy densities on the order of 1 GeV/fm^3 [\(11\)](#). These high energy densities exceed the theoretical minimum expected for Quark Gluon Plasma formation [\(5\)](#). The experimental capabilities of RHIC make it an invaluable tool in studying exotic forms of nuclear matter.

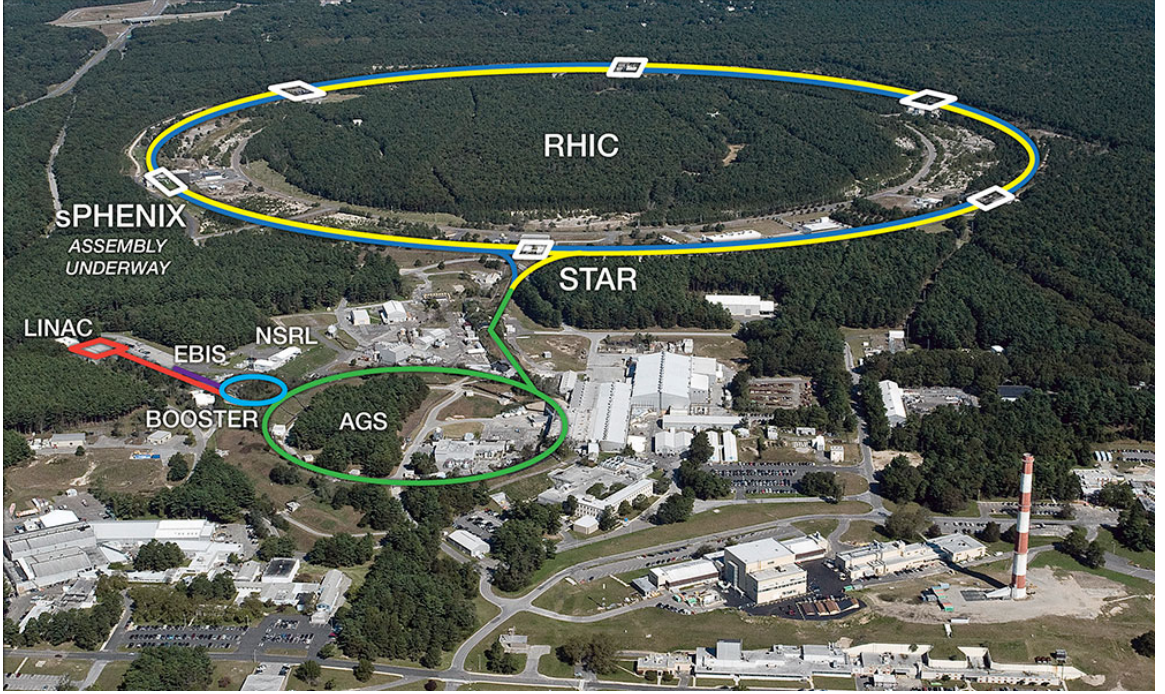


Figure 2.1: Aerial layout of RHIC
(12)

2.2 sPHENIX Detector

sPHENIX is the successor to the PHENIX detector that was completed at the Relativistic Heavy Ion Collider in April of 2023. sPHENIX uses both electromagnetic and hadronic calorimeters with full azimuthal acceptance, combined with Monolithic Active Pixel Sensor (MAPS) vertex and Time Projection Chamber (TPC) to facilitate a better analysis of heavy ion collisions (13). The detector was created to answer several questions regarding the QGP, such as the development and propagation of partonic showers in the material, investigations into asymptotically free theories on fundamental particles, and the details of transitions to QGP at the critical temperature, T_c (5).

The sPHENIX detector obtains data on the QGP by analyzing jets, collimated sprays of particles, with both electromagnetic and hadronic calorimetry. The physical design of the sPHENIX detector involves a 1.5T magnetic solenoid, tracking system,

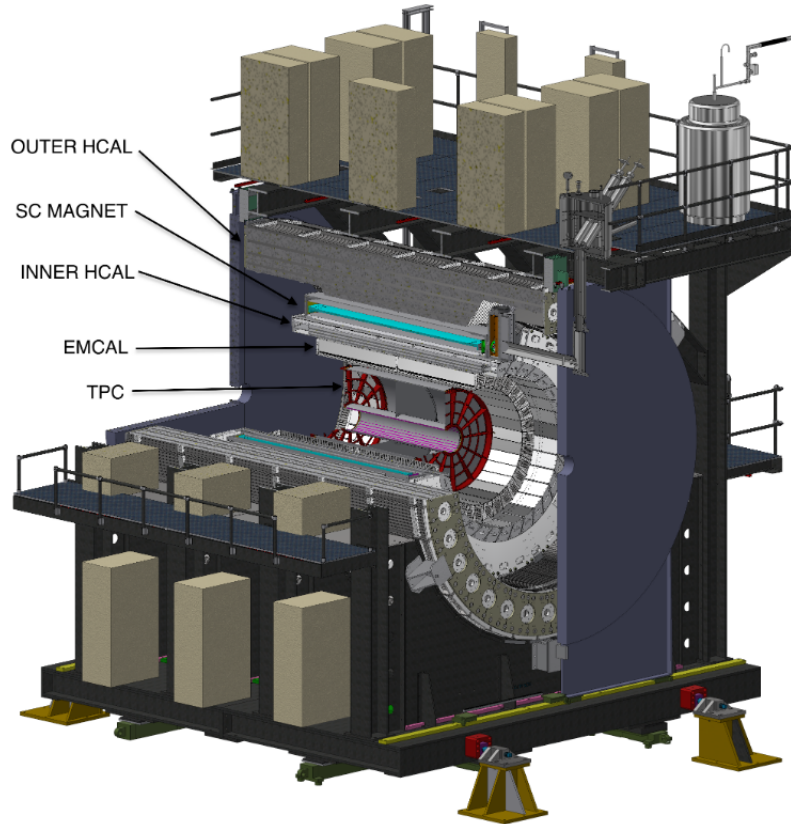


Figure 2.2: sPHENIX Detector and Components (1).

an electromagnetic calorimeter, and an inner and outer hadronic calorimeter. The measurement capabilities of the detector allows for the analyses of photons, hadrons, and jets, thus creating a useful experimental tool in addressing pressing questions regarding the dynamics of the QGP (1).

The data that sPHENIX records from jets are useful in constructing meaningful interpretations of the experiments that are conducted at the RHIC. Following a collision, ejected particles in jets are resolved by detectors, allowing them to be identified and properties such as their energy, momentum, and charge to be calculated.

2.3 Electromagnetic Calorimeter

The design, inner workings, and data collection capabilities of the EM calorimeter are an integral component to the understanding of the significance of the experiment, and the application of machine learning principles to the analysis of the data.

The EM calorimeter at sPHENIX is subject to a variety of constraints in regards to its design and performance, stemming both from the physical limitations imposed by its placement in sPHENIX, and from the physical attributes that it must be capable of measuring. In order to achieve sufficient accuracy in jet measurements, the EM calorimeter has a high degree of coverage for the solid angle encapsulated by sPHENIX, with the precise values for those angular constraints being denoted below in function 2.1 (14).

$$\begin{cases} |\eta| < 1.1 \\ \phi = 2\pi \end{cases} \quad (2.1)$$

It is of paramount importance that the calorimeter minimizes "dead zones" in its angular projections, as this will ensure good uniformity in measurements of electromagnetic observables (1).

What is perhaps the most important characteristic of the EM calorimeter is that it has the proper energy resolution to differentiate observables from background signatures. For the EM calorimeter, this translates to the ability to distinguish upilon signatures from irrelevant signatures present in the collider following an experimental trial. In relativistic heavy-ion collisions of energies approximately equal to 200GeV, this implies an energy resolution for the EM calorimeter of at least $15\%/\sqrt{E}$ (14). This is considered to be a precise enough resolution to properly identify observables generated in jet-inducing experiments, particularly photons, and the associated upilon states.

In addition to the requirements imposed on the EM calorimeter purely due to the nature of the physical quantities of interest, the detector is also subject to

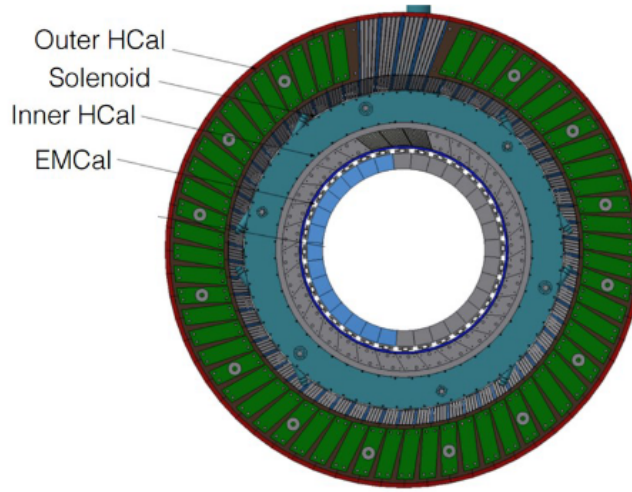


Figure 2.3: Cross Section of sPHENIX Calorimeters (1).

the constraints put in place by its position in the larger detector system, and its synchronization with other functionalities of sPHENIX. As can be seen above in figure 2.3, the EM calorimeter is situated inside of sPHENIX's magnetic solenoid. In order to achieve a compact size, and maintain functionality in the presence of strong magnetic fields, silicon photomultipliers/avalanche photodiodes have been used in the calorimeter in place of older photomultiplier tube technologies which are much less space efficient (5).

The EM calorimeter consists of repeating layers of a tungsten powder and epoxy mixture, interlaid with scintillating fibers (W/SciFi) (1). This arrangement of materials is configured into blocks that surround the detector in a cylindrical orientation, as shown by inspection of figure 2.4 below. The layers in the EM calorimeter include absorbers, and active detection layers. The tungsten/epoxy base material acts as an absorber for incident electromagnetic particles, such as the photons created by the charge-neutral pion decay. The interaction with the tungsten/epoxy material releases a shower of photons and electrons which permeate further into the detector. These electromagnetically interacting particles trigger light production in the scintillating fibers. The silicon photomultipliers will convert the

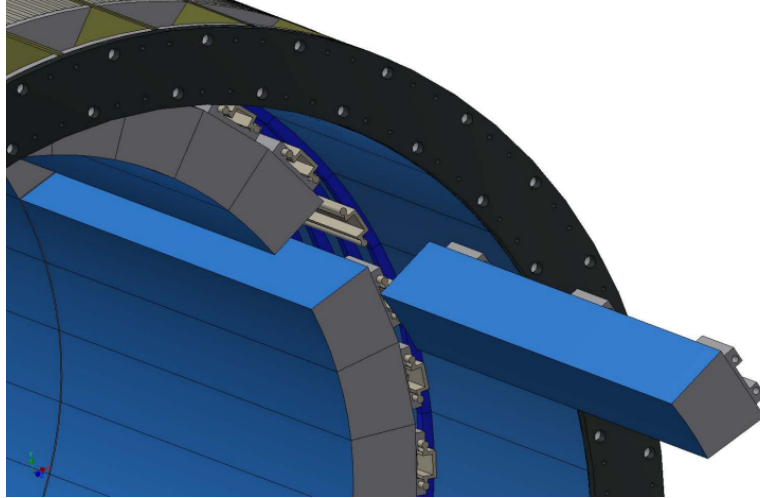


Figure 2.4: EM Calorimeter Blocks in sPHENIX (1).

light signatures from these particles into electrical signals which can be meaningfully interpreted by the detector readout (1).

For the EM calorimeter, the readout system is composed of 4 subsystems, including an analog signal system, controls, digitizers, and a power distribution system (1). After the signals are generated in the silicon photomultipliers, they are combined into a single channel, called the analog sum signal. This signal is then amplified, and then passed to the digitizer system.

Chapter 3

Analysis Method and Data

3.1 Analysis Goal

One method that has been explored in regards to improving the data analysis practices at sPHENIX is the implementation of machine learning to the classification of clustered energy signatures (15). Machine learning software is built upon the premise of computational pattern recognition from established samples, and the extrapolation of this pattern recognition onto new samples that have not been seen by the algorithm. With enough exposure and training to the material that is to be processed, these algorithms can learn to identify characteristics that associate a sample with a given class, type, etc., and apply this learned ability to new data. Machine learning has become more developed in recent years, and several sub-varieties of this software exist to facilitate the processing of several different types of data, ranging from image-based information to numerical representations.

3.2 Convolutional Neural Networks

One method that shows rising promise in recognizing these particle signatures on the detectors is the implementation of machine learning to the associated image

processing. Convolutional Neural Networks (CNN) are a class of deep learning-based machine learning algorithms that have become a quintessential tool in machine learning software where image processing is concerned. Convolutional Neural Networks are composed of layers of "neurons", aptly named for their resemblance to neural connectivity observed in the brains of living organisms. CNNs are composed of layers of these neurons, with varying types of connectivities used to filter, process, and extrapolate data. CNN's image processing capabilities make them a strong candidate to meet the needs of the sPHENIX experiment. Implementing a machine learning-based regime to the classification of particle signatures collected by the detectors would allow for a fast, accurate, and consistent methodology to be utilized (16).

3.3 Model Architecture

The model aims to alleviate the issues outlined in section 1.2 regarding the energy signatures collected by the EM calorimeter following neutral pion decay. The specific task that the model must complete is to classify the detector images as being either isolated or merged. Isolated images have energy signatures that are clearly separated, and merged images are those with signatures that lie very closely together.

The CNN model that is utilized in the project is created with Keras, which is a deep-learning API written in Python (17). It is built on top of Tensorflow, an open-source machine learning and AI library that is used for the creation and training of neural networks. The functionalities of TensorFlow are built upon the use of multi-dimensional arrays of data known as tensors. These tensors are connected with nodes between them, serving to represent mathematical functions which are performed on the arrays of data. These interconnected interfaces of nodes can be viewed on a more macroscopic scale as the layers that constitute the general makeup of Convolutional Neural Networks that was outlined in section 3.2. The pairing of Keras with the underlying architecture of TensorFlow was used to create the layers and functionalities of the model, the architecture of which can be seen above in figure 3.1.

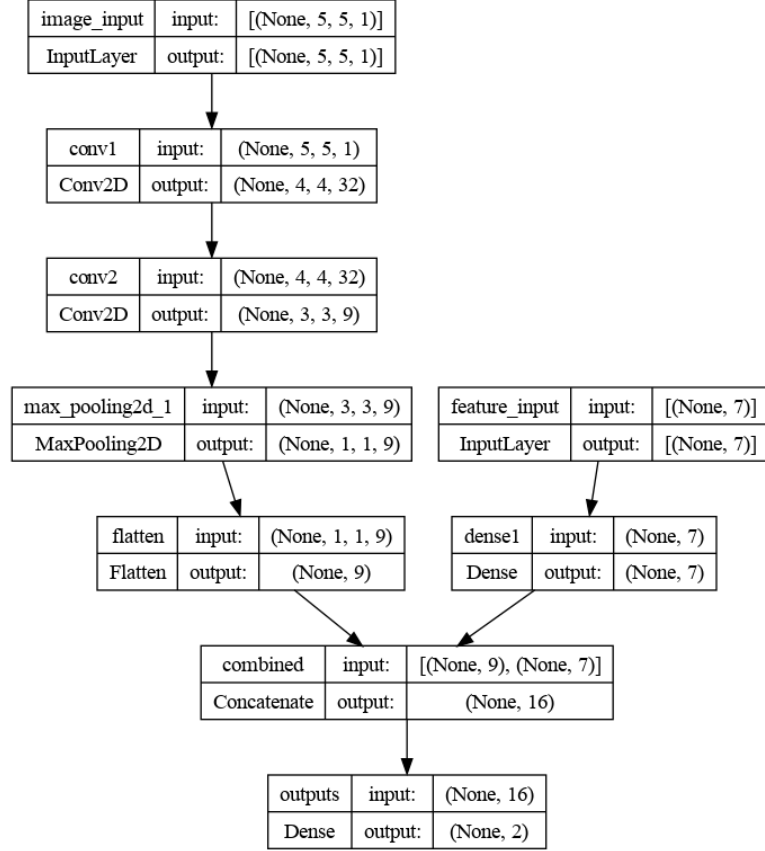


Figure 3.1: Model Architecture

As can be observed in the figure, several different components constitute the makeup of the machine learning model. The first item in the diagram is of the image input that the model receives. The image input is an array of the images collected from the EM calorimeter. They are characterized as heat maps of the photon energies. The model is trained on a combination of both merged and isolated image samples, some of which can be seen above in figure 3.2.

It can be seen in figure 3.1 that the machine learning model performs 2 convolutions on the image inputs immediately after receiving them. A convolution, in the context of machine learning, is designed to create feature maps for the input images. A feature map is simply a mapping of characteristics that the model learns to associate with a given outcome or class. The model creates feature maps by creating a 2-dimensional object known as a kernel, and superimposing it over an area of a

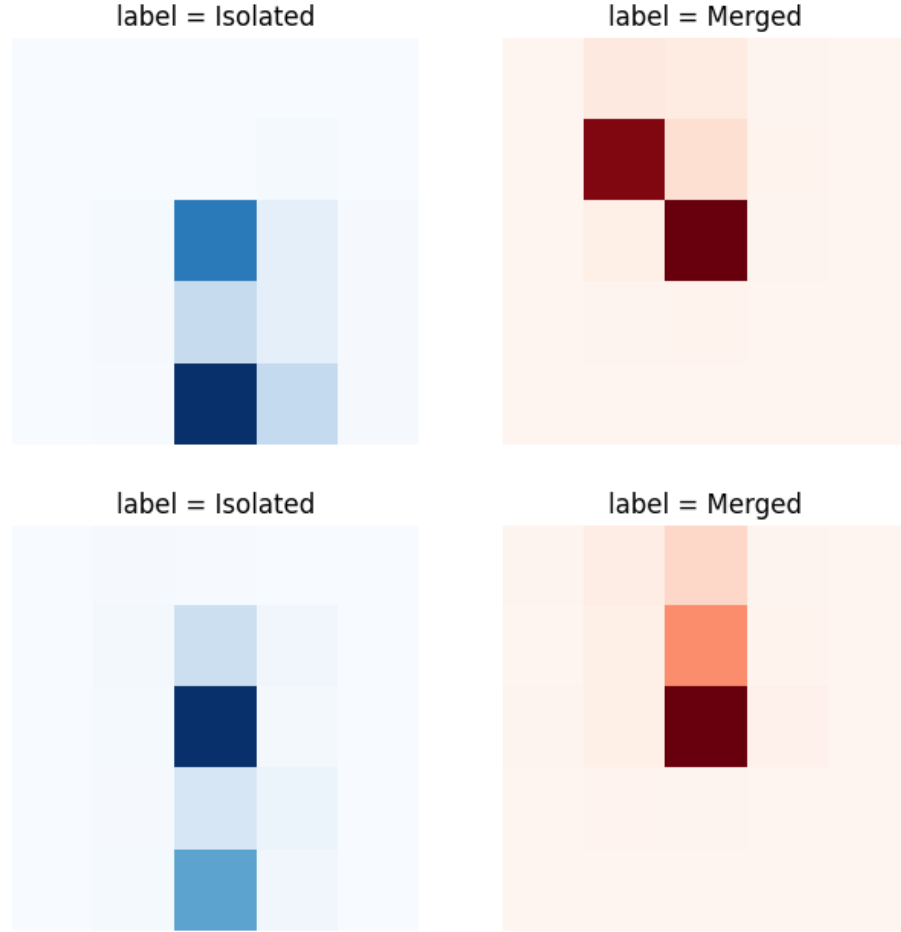


Figure 3.2: Merged and Isolated Sample Examples

given image. The kernel extracts features from the image, and places them onto a feature map. After it does this for a given area, it shifts to a new area to repeat this feature extraction. The kernel will repeat this process until it has iterated over the entire image, thus creating a feature map for each of the image inputs. (16).

Immediately following the convolution layers, the resultant feature map is passed to a pooling layer. The purpose of the pooling layer is to compress the feature map by selecting a point in the map, and using surrounding data in a given area to give a summarized representation of that area, thus reducing the complexity of the feature map. The model uses Max Pooling, which is a particular type of pooling operation where the function records only a maximum value within a given rectangular area

on the feature map. This allows for the map to be composed only of the most important data, without the unnecessary noise associated with weaker signatures. Pooling functions are especially useful because they make the feature maps translation invariant. Translational invariance dictates that the feature maps will convey the same information even if the locations of the features are shifted. In more macroscopic terms, it implies that the model will invest more of its learning into the presence (or the lack thereof) of certain features rather than focusing on where they are in a given image (16).

It is after the pooling function that another branch is incorporated into the model, indicated by the addition of the feature inputs. The feature inputs are composed of 5 different categories of data, including:

- Total energy in a cluster
- Azimuthal angle ϕ
- Pseudorapidity η
- Moment in xx, xy, and yy
- Probability of fit of a cluster

After incorporating the feature inputs to the model, they are passed to a dense layer. A dense layer, sometimes referred to as a fully connected layer, is named as such because the neurons that compose it are able to fully engage with those in layers occurring before and after their own. The dense layer operates by attaching a weight to each input that it receives from the input layer. Larger weight values indicate a more active connection between the neurons of the two layers. Dense layers typically incorporate the use of an activation function, in order to eliminate linearity in the data processing (16). The activation function that this model uses is ReLU, which stands for "rectified linear unit". When ReLU encounters an input value, it returns a

0 if that input happens to be negative. If the input value is positive, it simply returns that input value, unchanged (17).

After the feature inputs have been passed to the dense layer, the associated data from both the feature inputs and image inputs are combined, as can be seen in the concatenation operation in figure 3.1. This combination allows for the model to train on both types of inputs, and learn any associations between the two characteristics that may be present. The final output of the model comes from the resultant of this dataset being passed through a final dense layer. Instead of using the ReLU activation function like the previous dense layer, this one uses the softmax function. Softmax is utilized to associate the outputs of the dense layer with a probability distribution (17). This, in turn, allows the model to make predictions based on which class (merged or isolated) has the greatest probability for a given sample.

Chapter 4

Model Performance and Comparisons

4.1 ML Model Performance

The analysis of the performance of the machine learning model includes representations of the model training process, accuracy scores for both training and testing data, and a decomposition of the model's correct and incorrect predictions to demonstrate where the model is most and least successful in differentiating between the 2 sample classes.

Seen below in figure 4.1 is a depiction of the training process for the machine learning model. The graph legend describes 2 characteristics for both the training and testing phases. One such characteristic is the loss function. The purpose of the loss function is to express the magnitude of error between the model's predictions for the class of a given image and the true class. Naturally, the desired behavior for the loss function is for it to be minimized as the training progresses. As can be seen in the figure, the most rapid decrease in the loss function is observed in the first 20 training epochs. This is a normal tendency for most algorithms, as it becomes more difficult for the model to improve once a certain benchmark of accuracy is achieved.

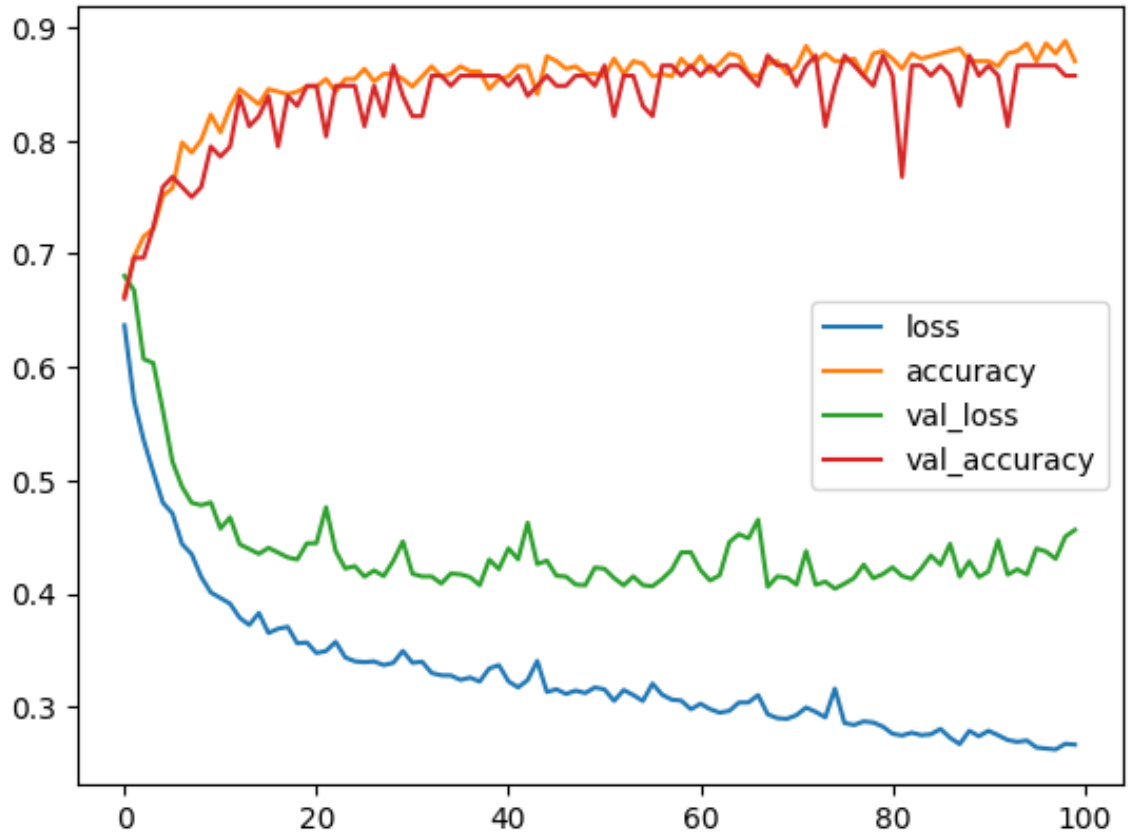


Figure 4.1: Algorithm Training Process Characteristics

This rapid descent and gradual plateau was present in both the training and testing loss functions, represented by the blue and green lines, respectively.

The accuracy trendlines simply denote the proportion of correct/incorrect predictions from the model. As can be seen in the figure, these trendlines increase rapidly within the first 20 training epochs, then reach an asymptote at approximately 90 percent accuracy. This accuracy can be further decomposed in figure 4.5, which is a confusion matrix representing the performance of the model on merged and isolated image samples.

One especially useful aspect of the confusion matrix is that it allows for trends in the model’s inaccuracies to be exposed. Instead of being presented with merely a numerical representation of the accuracy, the matrix shows the particular classes that the model’s prediction capabilities both excel and struggle with. To illustrate a qualitative analysis of the incorrectly predicted images, figures including both types of inaccuracies have been included below in figures 4.2 and 4.3:

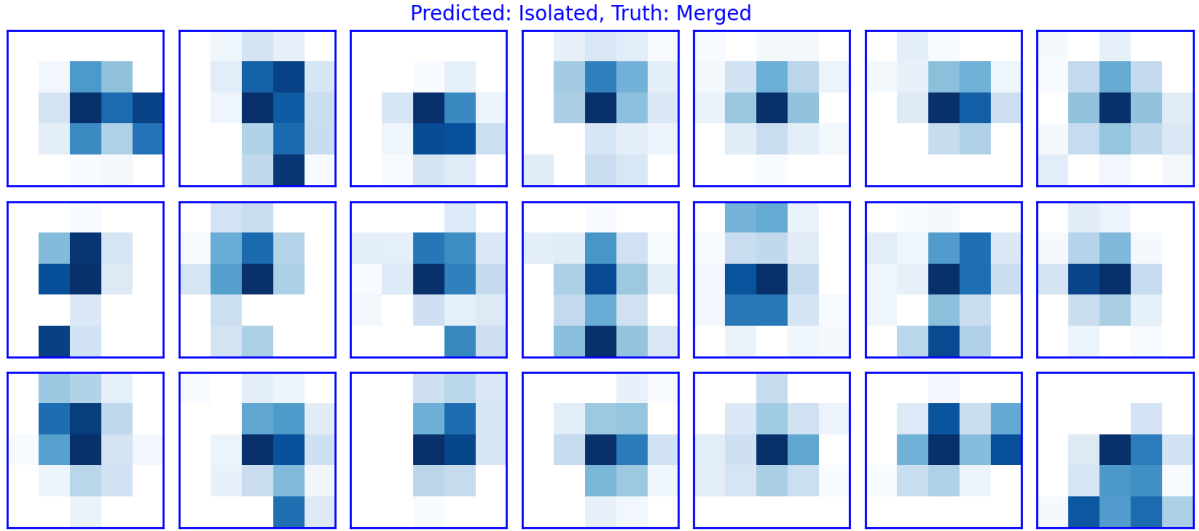


Figure 4.2: Incorrectly Predicted Merged Samples

Figure 4.2 above portrays the samples that are merged in actuality, but the model predicted as being isolated.

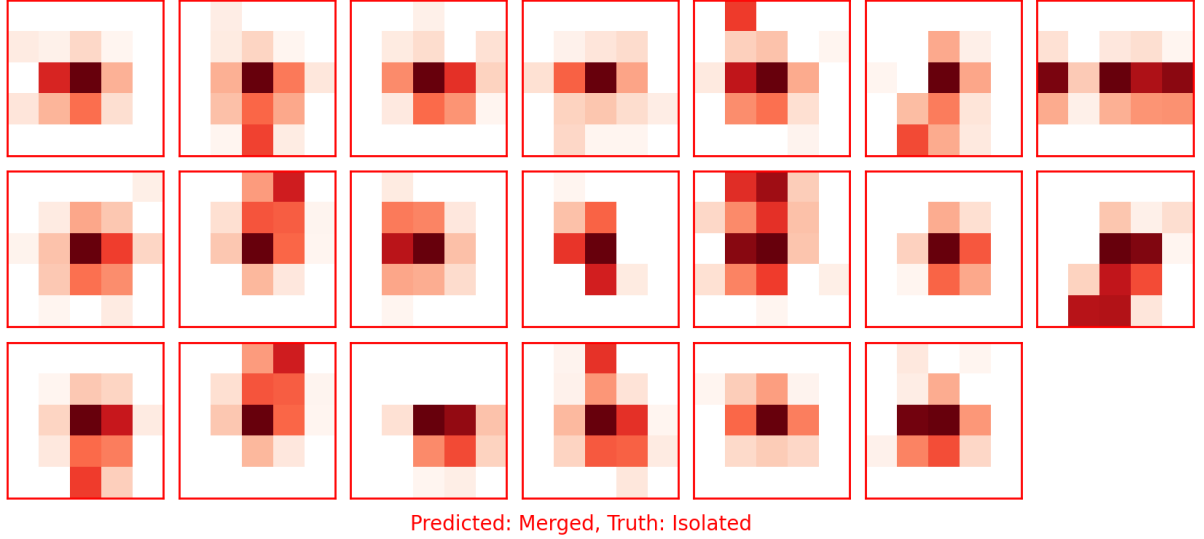


Figure 4.3: Incorrectly Predicted Isolated Samples

Figure 4.3 displays the inverse of the former error type. These samples are those which the model predicted to be merged when they are actually isolated.

In order to observe the differences between samples that are easily categorized by the model and those that are too ambiguous in their morphology for accurate predictions, figure 4.5 below has been included to show samples that the model predicted accurately. .

One prominent difference that can be observed between the correctly predicted samples and the incorrectly predicted samples is that the amount of whitespace in isolated vs merged samples is different. For the samples that the model predicted correctly, isolated images have approximately 16 percent more whitespace than merged images. In the incorrectly predicted dataset, the merged samples have approximately 19 percent more whitespace than isolated samples. It is very possible that through the model's training, an association was made between large amounts of whitespace and likelihood of being an isolated sample. Because the incorrect samples go against this common trend, this could be a reason why they were not predicted with accuracy.

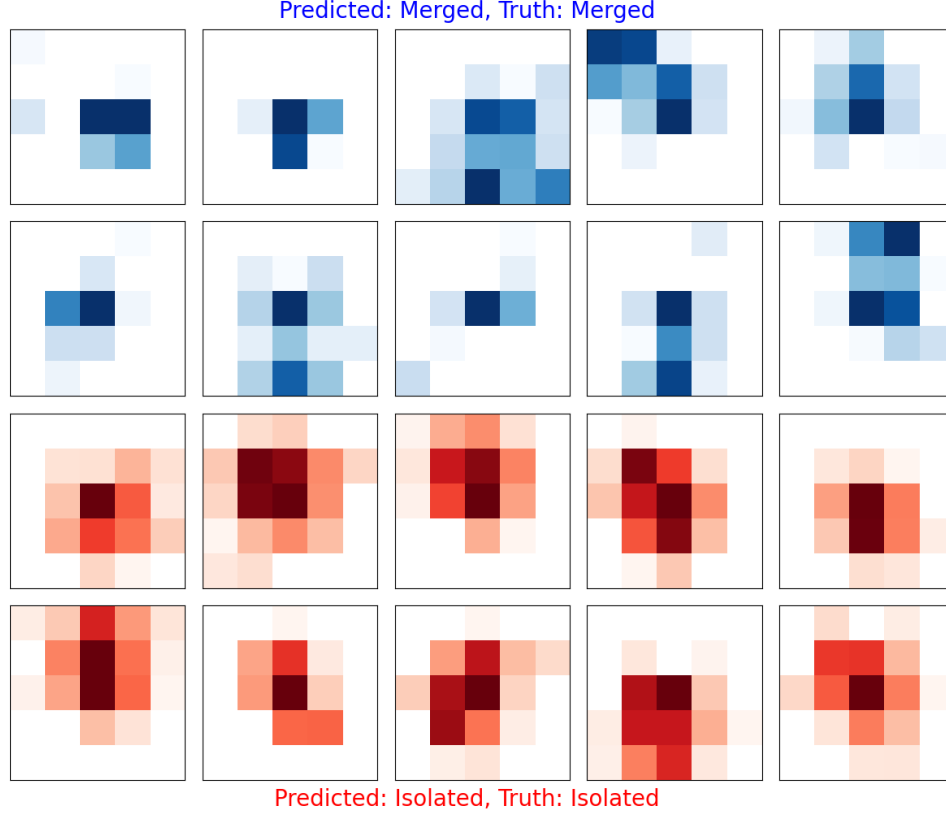


Figure 4.4: Correctly Predicted Samples

4.2 Comparisons of Performance

In comparison to the traditional method of classification, the machine learning model is considerably more accurate when handling both isolated and merged samples. In the figure below, confusion matrices representing the performance of both the traditional method and the machine learning model can be found. For merged samples, the machine learning model demonstrates 9.86% more accuracy than the traditional method. For isolated samples, the accuracy of the machine learning model is 20.52% better than the traditional method. In regards to incorrect predictions, both the machine learning model and the traditional method demonstrate the most inaccuracy with merged samples, with merged samples making up 37.13% and 27.27% of inaccuracies for the traditional method and machine learning model, respectively.

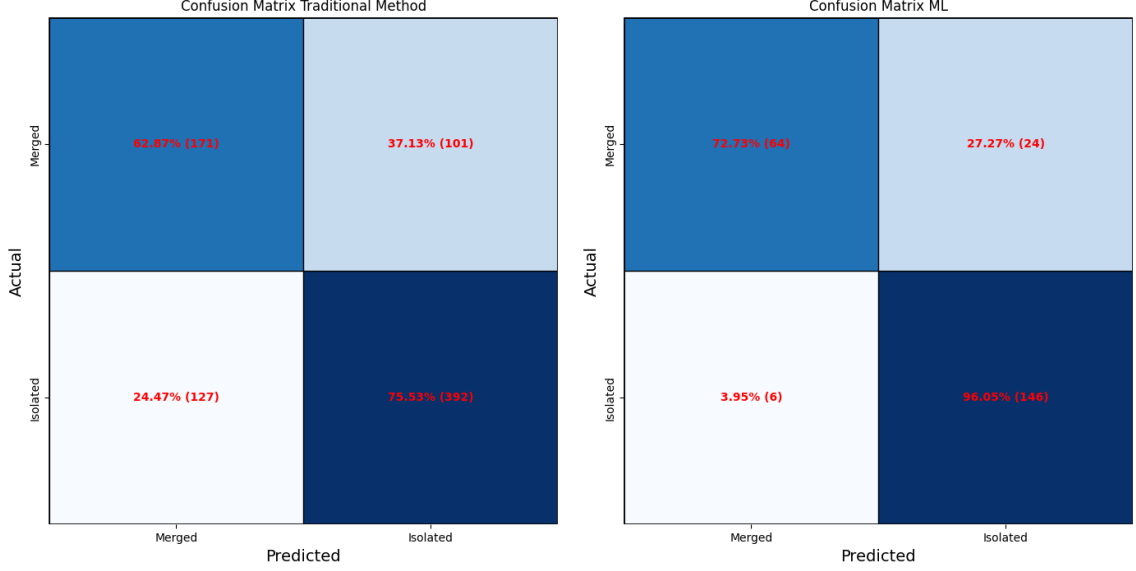


Figure 4.5: Confusion Matrix for Traditional Method (Left) compared to the confusion Matrix for ML Model (Right).

Isolated samples propagate fewer errors in both. The traditional method exhibited 24.47% error with this class of sample, and the machine learning model 3.95%.

The plot below in figure 4.6 depicts a Receiver Operating Characteristic (ROC) curve for the two methods of image classification. The vertical axis measures the true positive rate. The true positive rate is numerically calculated from the formula:

$$TPR = \frac{\text{True Positives}}{\text{True Positives} + \text{False negatives}}$$

Where True Positives are defined as correctly identified samples of a given class. False negatives are samples which are of that same class which are not properly identified. The true positive rate is therefore the proportion of correctly identified samples of a given class out of the total number of samples belonging to that class. The horizontal axis shows the false positive rate, which is denoted by:

$$FPR = \frac{\text{False Positives}}{\text{False Positives} + \text{True negatives}}$$

The false positive rate is therefore the proportion of samples incorrectly identified as being of a given class to the total number of samples that do not belong to that class.

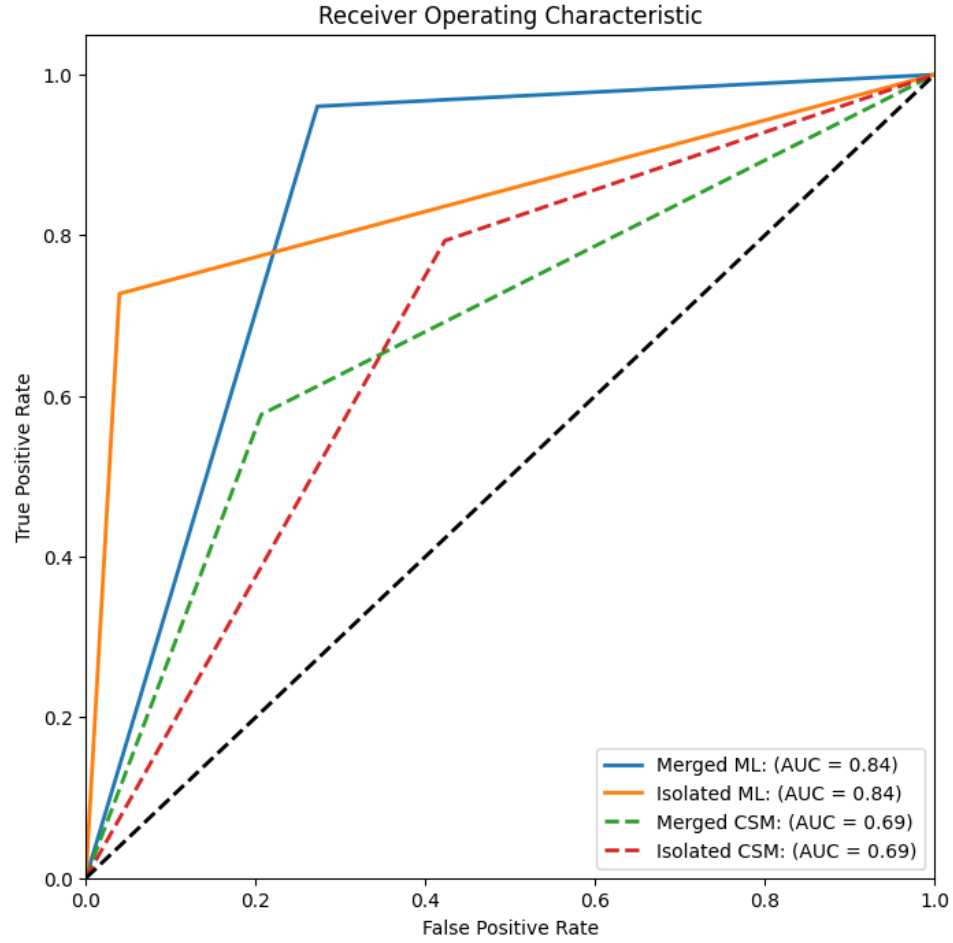


Figure 4.6: Receiver Operating Characteristic (ROC) Curve

It can be seen in the plot that the machine learning model reaches higher values for the true positive rate than the traditional method. Also, the machine learning model does not reach as high of values for the false positive rate as the traditional method. This behavior is indicative of substantially better accuracy for the machine learning model, as it is more likely to correctly identify samples of a given class, with fewer false positives.

Chapter 5

Discussion

5.1 Energy and Whitespace Correlations

One of the most physically meaningful trends observed in the analysis of the machine learning model's performance is the impact that image whitespace has on the accuracy of the predictions. As was found in section 4.1, the proportions of whitespace for isolated and merged samples is reversed for incorrectly predicted and correctly predicted samples. For the correctly predicted samples, isolated images have much more whitespace, averaging approximately 16 percent more than merged samples. For incorrectly predicted samples, merged images have approximately 19 percent more whitespace than isolated images.

In order to establish the physical significance of this difference, it is important to recall the means by which the signatures on the images are created. As was outlined in section 1.2, the signatures on the images are representative of the energy imparted to the EM detector by photons created in neutral pion decay. Nonzero (shaded) portions of the heat maps therefore represent areas with given energy values associated with them. Areas with low energy measurements correspond to lighter shades, which darken as energy measurements increase beyond certain threshold values. Whitespace

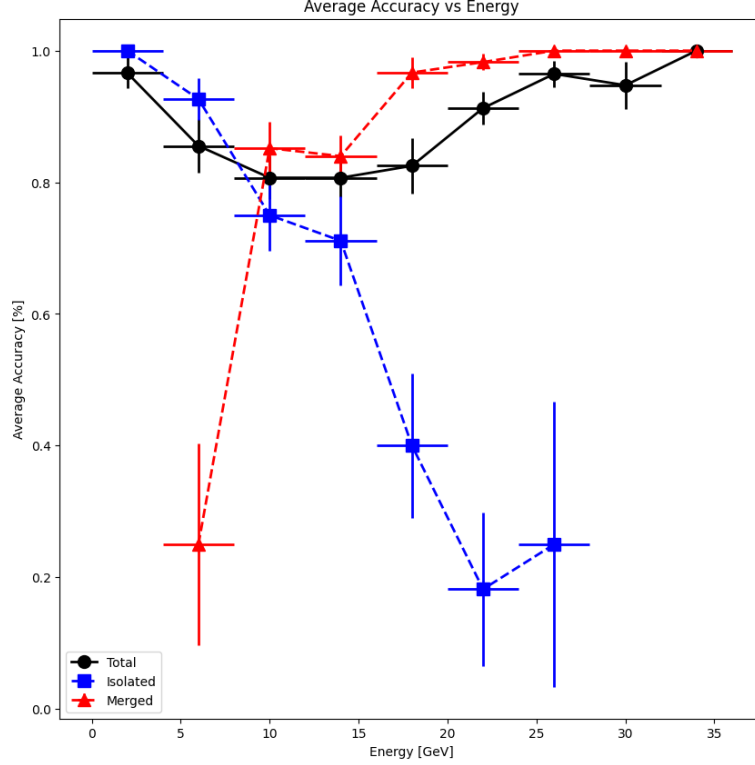


Figure 5.1: Accuracy of ML Model as a Function of Sample Energy

is present in areas where energy measurements are below a certain threshold value which has been set to selectively ignore noise, or insignificant energy levels.

Seeing that the energies of incident photons contribute heavily to the morphology of the images, it is reasonable to assume that the model would correlate the energies associated with a given sample to its likelihood of being isolated or merged. The accuracy of the model at different energy levels can be found above in figure 5.1.

It is quite evident from analyzing the accuracy of the model as a function of energy that the model does make an association between the energy of a sample and whether it is likely to be merged or isolated. For merged images, the accuracy is very low for samples with energies below approximately 10 GeV, showing that the model makes a strong association between high energy samples and the likelihood of them being merged. For isolated samples, the opposite is observed. The effectiveness of the

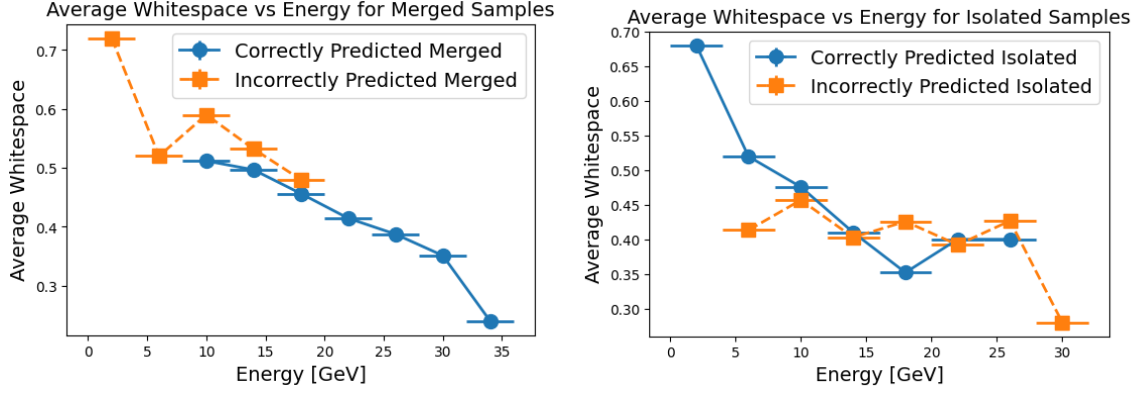


Figure 5.2: Average Whitespace vs Energy for Merged Samples (Left) and Average Whitespace vs Energy for Isolated Samples (Right).

model at predicting isolated samples decreases steadily as the energies of the samples go up, showing that the model associates isolated images with lower energies.

Seeing that the energies of an image correlate with the whitespace that is present within it, figure 5.2 above represents the relationship between whitespace and energy for both correctly and incorrectly merged and isolated samples.

It can be seen in the left image that accuracy for the merged samples is the greatest in the $> 20\text{GeV}$ range, as the number of incorrectly predicted samples goes down, while the number of correctly predicted samples maintains a relatively high value. We can see that in this region of high merged sample accuracy that the average whitespace is lower than the 10-20GeV region where the accuracy isn't as high.

In the image on the right, it can be seen that the isolated images are predicted most accurately in the $< 10\text{GeV}$ range, denoted by the high number of correctly identified samples compared to incorrectly identified samples. This region of high accuracy corresponds to very high average whitespace values, at nearly 70 percent.

5.2 Physical Implications

From the analysis of section 5.1, meaningful connections between the performance of the model and the physical properties of the pion decay can be made. It is quite evident that the ML model regards the energy and amount of whitespace of a given sample as important metrics in sorting between merged and isolated samples. It is apparent that high energy, low-whitespace samples are strong indicators of an image being that of narrowly-separated photons from high-momentum pion decay, which characterizes a merged image. Low energy, high-whitespace samples are more characteristic of isolated samples, which are produced from low-momentum pion decay. This makes logical sense, as the photons produced from low momentum decay will have less energy imparted to them, and thus more whitespace. High momentum decay will create more energetic photons, and thus less whitespace.

Cases in which this tendency is defied, such as with the incorrectly predicted samples, are illuminating in regards to ways that the EM calorimeter could be improved. For incorrectly predicted merged samples, which represent the majority of ML model inaccuracies, it is likely that these samples are of lower energy, with higher amounts of whitespace associated with their heat map signatures. As mentioned before, the heat maps signatures are generated based on threshold values that must be met to register a nonzero signature in a given area. It is very possible that small energy contributions from the merged photons toward the periphery of the images are being truncated by the detector. This truncation is a function of the energy resolution of the detector. Improving the energy resolution in the detector would allow for less truncation to take place, thus making the images more inclusive of small energy signatures. It is evident through the analysis of the performance of both the machine learning model and the traditional method that merged images are the harder of the 2 classes for both to sort. If less whitespace was removed from these images, then they would more closely resemble their low-whitespace archetype, making them more recognizable to the sorting algorithms.

Bibliography

- [1] sPHENIX collaboration, “sPHENIX Technical Design Report,” 2019. [ix](#), [6](#), [7](#), [8](#), [9](#)
- [2] W. Busza, K. Rajagopal, and W. van der Schee, “Heavy ion collisions: The big picture and the big questions,” *Annual Review of Nuclear and Particle Science*, vol. 68, p. 339–376, Oct. 2018. [1](#), [2](#)
- [3] Y. Akiba *et al.*, “The Hot QCD White Paper: Exploring the Phases of QCD at RHIC and the LHC,” 2 2015. [1](#)
- [4] J. N. Guenther, “Overview of the QCD phase diagram: Recent progress from the lattice,” *Eur. Phys. J. A*, vol. 57, no. 4, p. 136, 2021. [1](#)
- [5] A. Adare *et al.*, “An Upgrade Proposal from the PHENIX Collaboration,” 1 2015. [1](#), [4](#), [5](#), [8](#)
- [6] G. David, E. Kistenev, S. White, C. Woody, A. Bazilevsky, S. Belikov, V. Kochetkov, V. Onuchin, and A. Usachev, “Pattern recognition in the phenix pbsc electromagnetic calorimeter,” in *1998 IEEE Nuclear Science Symposium Conference Record. 1998 IEEE Nuclear Science Symposium and Medical Imaging Conference (Cat. No.98CH36255)*, vol. 1, pp. 520–524 vol.1, 1998. [3](#)
- [7] K. Adcox *et al.*, “Formation of dense partonic matter in relativistic nucleus nucleus collisions at RHIC: Experimental evaluation by the PHENIX collaboration,” *Nucl. Phys.*, vol. A757, pp. 184–283, 2005. [4](#)

- [8] J. Adams *et al.*, “Experimental and theoretical challenges in the search for the quark gluon plasma: The STAR collaboration’s critical assessment of the evidence from RHIC collisions,” *Nucl. Phys.*, vol. A757, pp. 102–183, 2005. [4](#)
- [9] B. B. Back *et al.*, “The PHOBOS perspective on discoveries at RHIC,” *Nucl. Phys.*, vol. A757, pp. 28–101, 2005. [4](#)
- [10] I. Arsene *et al.*, “Quark Gluon Plasma an Color Glass Condensate at RHIC? The perspective from the BRAHMS experiment,” *Nucl. Phys.*, vol. A757, pp. 1–27, 2005. [4](#)
- [11] A. Bazavov, F. Karsch, Y. Maezawa, S. Mukherjee, and P. Petreczky, “In-medium meson properties and screening correlators,” *Journal of Physics: Conference Series*, vol. 535, p. 012031, sep 2014. [4](#)
- [12] B. <https://www.bnl.gov/newsroom/news.php?a=119262>, “Bnl news letter,” 2021. [5](#)
- [13] C. A. Aidala *et al.*, “Design and Beam Test Results for the sPHENIX Electromagnetic and Hadronic Calorimeter Prototypes,” *IEEE Trans. Nucl. Sci.*, vol. 65, no. 12, pp. 2901–2919, 2018. [5](#)
- [14] C. A. Aidala, V. Bailey, S. Beckman, R. Belmont, C. Biggs, J. Blackburn, S. Boose, M. Chiu, M. Connors, E. Desmond, A. Franz, J. S. Haggerty, X. He, M. M. Higdson, J. Huang, K. Kauder, E. Kistenev, J. LaBounty, J. G. Lajoie, M. Lenz, W. Lenz, S. Li, V. R. Loggins, E. J. Mannel, T. Majoros, M. P. McCumber, J. L. Nagle, M. Phipps, C. Pinkenburg, S. Polizzo, C. Pontieri, M. L. Purschke, J. Putschke, M. Sarsour, T. Rinn, R. Ruggiero, A. Sen, A. M. Sickles, M. J. Skoby, J. Smiga, P. Sobel, P. W. Stankus, S. Stoll, A. Sukhanov, E. Thorsland, F. Toldo, R. S. Towell, B. Ujvari, S. Vazquez-Carson, and C. L. Woody, “Design and beam test results for the sphenix electromagnetic

- and hadronic calorimeter prototypes,” *IEEE Transactions on Nuclear Science*, vol. 65, p. 2901–2919, Dec. 2018. 7
- [15] A. Tondryk and S. Lee, “Aanalysis of Using Machine Learning for the Classification of Particles in Jets,” 2020. 10
- [16] I. J. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016. <http://www.deeplearningbook.org>. 11, 13, 14
- [17] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, “TensorFlow: Large-scale machine learning on heterogeneous systems,” 2015. Software available from tensorflow.org. 11, 15