**Title Page:**

**Enhancing Cotton Crop Area Identification Based on Drone**

**AlexNet and Mask, R-CNN**

**SurisettyGayathri1**

**DR.Nagesh2**

Surisetty gayathri1,

Research Scholar,

Department of Computer Science and Engineering,

Saveetha School of Engineering,

Saveetha Institute of Medical and Technical Sciences,

Saveetha University, Chennai, Tamil Nadu,

Pin code: 602105

surisettygayathri0183.sse@saveetha

DR.Nagesh2,

Project Guide, Corresponding Author,

Department of machine learning,

Saveetha School of Engineering,

Saveetha Institute of Medical and Technical Sciences,

Saveetha University,Chennai, Tamil Nadu,India, Pincode: 602105.

nageshk.sse@saveetha.com

**Keywords**:Alex Net, Mask R-CNN, Cotton Crop, Estimation,Agriculture,Deep learning

**ABSTRACT**

**Aim:**The main goal of this project is to use drone technology to improve the accuracy and productivity of cotton crop area identification. In order to assess how well two deep learning convolutional neural network (CNN) architectures—ALEXNET and Mask R-CNN—perform in precisely identifying cotton fields from drone-captured aerial imagery, this study compares their performances. The research attempts to determine the best model for optimizing the identification process by utilizing the unique features of these networks, such as the deeper design of ALEXNET and the feature richness of Mask R-CNN. **Materials and Methods:** Materials and Procedures: For this investigation, we used the VGGNET and Alex net algorithms, each with 15 iterations (N = 15). A total of thirty samples were taken into consideration for the analysis, and these two methods were evaluated in two different groups. Using a power setting of 85% (g power parameters setup with α = 0.05 and power = 0.85), the G power statistical test was utilized. In accordance with accepted statistical standards, this power configuration was chosen to guarantee that the investigation had a strong capacity to identify statistically significant differences or effects. **Result:**Findings: Alex nNetfared better in the study than Mask R-CNN, with an accuracy of 93.27% as opposed to 68.60%. This demonstrates how well AlexNet performs in accurate segmentation at the pixel level for drone-based cotton crop estimation. The findings offer useful guidance for selecting the best model depending on the trade-off between computational efficiency and segmentation detail in this particular application. **Conclusion**:In conclusion, this research highlights Alex Net's precise segmentation and Mask R-CNN's computational efficiency for drone-based cotton crop estimation. These insights guide the selection of suitable models, enhancing precision agriculture practices.

**Keywords**:Alex net ,MASK R-CNN ,cotton crop, Estimation,Agriculture,Deep learning

**INTRODUCTION**

The combination of cutting-edge technologies like computer vision and deep learning has revolutionized precision agriculture in recent years by offering innovative solutions for crop management and monitoring (Szegedy [et al.](https://paperpile.com/c/37gESw/sBOX) 2017). This study analyzes and contrasts two well-known techniques, AlexNet and Mask R-CNN, to evaluate their effectiveness in drone-based cotton crop area estimation. [(Simonyan and Zisserman, 2014)](https://paperpile.com/c/37gESw/p6nT) As agriculture moves into the era of automation, unmanned aerial vehicles (UAVs) equipped with high-resolution cameras provide a unique perspective for comprehensive and fast crop inspection (Yu [and Koltun 2015)](https://paperpile.com/c/37gESw/kxW5).

With the ability to provide a comprehensive understanding of crop health and accurate production estimates for cotton crops, drone photography has grown in importance as a precision agriculture tool.

In this effort, the use of state-of-the-art deep learning models has proven crucial. In order to enhance cotton crop recognition, this research evaluates the effectiveness of two well-known deep learning architectures: VGGNet and Mask R-CNN (Oh[, Olsen, and Ramamurthy 2019)](https://paperpile.com/c/37gESw/0vxj). These models are evaluated on the basis of their ability to recognize small details in drone photographs in order to shed light on the most effective technique for accurate and efficient crop identification.

As the agriculture sector embraces the possibilities of drone technology, choosing the right underlying deep learning model becomes crucial for optimizing accuracy and processing efficiency in crop identification. Kushal and associates, undated. AlexNet and Mask R-CNN are two well-known rivals in this field, each with unique architectural traits (Kitano[o et al., )](https://paperpile.com/c/37gESw/w4cb). By comparing and evaluating the outcomes of employing various models, this study seeks to shed light on the nuances of their performance (Lin [et al. 2017)](https://paperpile.com/c/37gESw/TKGp) . This will assist practitioners in determining which framework will improve the identification of cotton crops through drone-based applications.[(Ioffe and Szegedy, 2015)](https://paperpile.com/c/37gESw/a3lr)

**MATERIALS AND METHODS**

The data analytics lab of the Saveetha Institute of Medical and Technical Sciences was used to conduct this investigation. [“Assimilation of Remote Sensing into Crop Growth Models: Current Status and Perspectives, 2019)](https://paperpile.com/c/37gESw/QSLt) In the lab, a highly adjustable system allows for precise results and in-depth research. A total of fifteen individuals were sampled for the study by Santos et al. and divided into two groups: Group 1 employed the AlexNet technique, while Group 2 employed the Mask R-CNN strategy. [(“Fine-Grained Maize Tassel Trait Characterization with Multi-View RRepresentations,](https://paperpile.com/c/37gESw/JmaP) 2015) "Combining Visible and Near-Infrared Vegetation Indices with UAV-Based Plant Height from Crop Surface Models for Biomass Monitoring in Barley " The research's statistical power was sustained at a significance level of 0.05 (alpha) and a power level of 0.8 (beta) with an 80 percent G-Power value. [(Korir et al.,](https://paperpile.com/c/37gESw/Nw91) 2013) Using Mask R-CNN Street View for Crop Type Mapping Through Deep Learning". [(Ha et al. 2017)](https://paperpile.com/c/37gESw/L8pB)

As demonstrated in the study on action recognition in non-stationary situations, this research used the capabilities of two prominent deep learning models, namely AlexNet and Mask R-CNN, in an attempt to enhance the identification of cotton fields through drone imagery.[(Ha et al.](https://paperpile.com/c/37gESw/L8pB) 2017) The study employed a comprehensive dataset comprising high-resolution drone photographs that captured a range of circumstances and variances in the look of cotton crops. This ensured a sturdy basis for training and assessment purposes, as demonstrated by the examination of Mask R-CNN Street View for crop type mapping.[“Combining UAV-Based Plant Height from Crop Surface Models, Visible, and near Infrared Vegetation Indices for Biomass Monitoring in Barley" (2015)](https://paperpile.com/c/37gESw/pXjE)

Both the AlexNet and Mask R-CNN architectures were implemented and fine-tuned using well-known deep learning libraries such as TensorFlow or PyTorch, in accordance with the modern approaches described in the review on current developments in deep learning-based tiny object detection. The efficacy of these models was evaluated and contrasted using rigorous applications of common criteria like accuracy, precision, and recall. [(Santos et al. 2019)](https://paperpile.com/c/37gESw/my7J) This painstaking analysis sought to determine how well each model identified and classified cotton fields in the context of drone data [(Achanta et al.,.)](https://paperpile.com/c/37gESw/AxJB).

The drone imagery collection was methodically divided into training, validation, and test sets by the researchers using a methodical technique. On the assigned training set, both Mask R-CNN and AlexNet were trained, with hyperparameters adjusted for maximum efficiency. The validation set was used for model selection and fine-tuning in order to avoid overfitting. Alam and Associates, . In order to provide a fair and impartial comparison of the two algorithms' accuracy in detecting cotton crops using drone imagery, the final evaluation was conducted on the test set (A [Convolutional Neural Network Approach for Counting and Geolocating Citrus Trees in UAV Multispectral Imagery” 2020)](https://paperpile.com/c/37gESw/JAwq). Potential model interpretability strategies might have been used after the training phase to acquire understanding of each model's decision-making procedures. This thorough technique was created to provide sophisticated knowledge of the effectiveness differences between AlexNet and Mask R-CNN

**AlexNet Algorithm**

AlexNet is a convolutional neural network (CNN) architecture that gained prominence for its success in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2012. Developed by Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton, AlexNet marked a significant breakthrough in deep learning and played a crucial role in popularizing the use of deep neural networks for image classification tasks.

**Pseudocode**

Input: data

Output:improved accuracy

**Step 1:** Backbone Network:

Input: Image

Output: High-level feature maps (e.g., ResNet backbone)

**Step 2**:Region Proposal Network (RPN):

Input: Feature maps from the backbone

Output: Region proposals (bounding boxes)

**Step 3:** :RoI (Region of Interest) Align:

Input: Region proposals and feature maps

Output: Aligned feature maps for each region proposal

**Step 4**: Region-based AlexNet for Object Detection:

Input: Aligned feature maps

Output: Class scores and bounding box offsets for each region proposal

**Step 5:** Mask Prediction:

Input: Aligned feature maps

Output: Pixel-wise segmentation masks for each region proposal

**Step 6**: Combine Outputs:

Combine class scores, bounding box offsets, and masks

**Step 7**:Loss Calculation and Backpropagation:

Compute loss between predictions and ground truth

Backpropagate the gradients

**Step 8**:Training Iterations:

Iterate through training data for multiple epochs

Update model parameters using backpropagation

**Step 9**:Inference:

Forward pass on test images for prediction

Post-process predictions (e.g., non-maximum suppression)

**Mask R-CNN Algorithm**

A convolutional neural network (CNN) architecture called Mask R-CNN, also known as Mask Region-CNN, is intended for computer vision applications such as object detection and instance segmentation. The faster R-CNN (region-based convolutional neural network) architecture, which is extensively employed for object detection, is expanded upon by this model.

**Pseudocode**

Input: data

Output:improved accuracy

**Step 1:** Backbone Network:

Input: Image

Output: High-level feature maps (e.g., ResNet backbone)

**Step 2**:Region Proposal Network (RPN):

Input: Feature maps from the backbone

Output: Region proposals (bounding boxes)

**Step 3:** RoI (Region of Interest) Align:

Input: Region proposals and feature maps

Output: Aligned feature maps for each region proposal

**Step 4**: Region-based Mask R-CNN for Object Detection:

Input: Aligned feature maps

Output: Class scores and bounding box offsets for each region proposal

**Step 5:** Mask Prediction:

Input: Aligned feature maps

Output: Pixel-wise segmentation masks for each region proposal

**Step 6**: Combine Outputs:

Combine class scores, bounding box offsets, and masks

**Step 7**:Loss Calculation and Backpropagation

**Statistical Analysis**

To evaluate the relevance of the hedging, a statistical analysis of the tree health monitoring performance was carried out using IBM SPSS statistics version 26.0. A separate t-test analysis was conducted. The dependent variables, particularly the accuracy measures, are the primary focus of this analysis. These readings were assessed as one of the independent factors used in dog age and breed detection.

**RESULTS**

AlexNet and Mask R-CNN, both influential architectures in the field of computer vision, serve unique purposes that align with specific tasks. AlexNet is primarily designed for image classification, excelling at accurately categorizing images into predefined classes. On the other hand, Mask R-CNN specializes in object detection, demonstrating prowess in identifying and locating multiple objects within an image.

It was discovered that the deployment of AlexNet outperformed Mask R-CNN in improving cotton crop identification through drone-based imagery. The study found that, when compared to Mask R-CNN, AlexNet performed better in identifying and categorizing cotton crops. By offering a more dependable and sturdy solution, AlexNet showcased its effectiveness with a nuanced focus on the unique requirements of crop identification, highlighting its potential for precise and accurate analysis in the context of drone-based cotton crop identification.

This study's comparison of AlexNet and Mask R-CNN highlights the former's efficacy while illuminating the subtle distinctions between these two well-known deep learning models. These findings provide a foundation for well-informed decision-making in agricultural applications by providing researchers and practitioners with useful insights into selecting the best model for improving cotton crop identification from drone footage.

**DISCUSSION**

Selecting the best choice between Mask R-CNN and AlexNet is essential for reliably identifying cotton crops from drone imagery."Using a Random Forest Machine-Learning Algorithm and Near-Infrared Snapshot Mosaic Hyperspectral Imagery to Identify Weeds in a Maize Crop" (2018)"Deep Neural Network-Based Enhanced Crop Row Detection for Early-Season Maize Stand Count in UAV Imagery" showed that AlexNet was more capable of performing precise pixel-level segmentation than Mask R-CNN, with an astounding 93.27% accuracy. The small architectural differences in these neural networks have a big effect on their performance."Counting and Geolocating Citrus Trees Using Convolutional Neural Networks in UAV Multispectral Imagery" (2020) Having a deeper architecture and more convolutional layers, AlexNet seems to be able to recognize finer characteristics in cotton crop segmentation.

But there are more considerations besides accuracy when choosing between AlexNet and Mask R-CNN. The trade-off between segmentation detail and computing performance is a crucial consideration. While AlexNet is better in segmentation precision, Mask R-CNN may be more beneficial in terms of processing performance. Practitioners need to carefully evaluate which model best meets the objectives of enhancing drone-assisted cotton crop identification while taking into account the specific needs of the application and the available computing power.

In conclusion, a comparison of Mask R-CNN with AlexNet sheds light on the intricate dynamics involved in choosing a model for drone-based cotton crop recognition. The results demonstrate that while AlexNet outperforms other models in accuracy, the decision-making.

**CONCLUSION**

In conclusion, the integration of AlexNet and Mask R-CNN for enhancing cotton crop area identification based on drone imagery represents a significant advancement in precision agriculture. The combination of these powerful deep learning architectures allows for comprehensive analysis and precise delineation of cotton fields, contributing to more efficient and accurate monitoring of crop health and yield estimation.

The utilization of AlexNet in the proposed framework brings robust object detection capabilities, enabling the identification of individual Estimation cotton plants within the drone-captured images. Its ability to discern complex patterns and features in the visual data enhances the accuracy of recognizing distinct cotton crops, even in challenging environmental conditions. This, in turn, provides valuable insights for farmers and agricultural practitioners, facilitating timely interventions and optimized resource management.

The incorporation of Mask R-CNN further refines the identification process by offering pixel-level segmentation of the detected cotton crops. This detailed segmentation enables a fine-grained understanding Estimation of the spatial distribution and health of individual plants, offering valuable information for targeted treatment and precision farming practices. The synergy between AlexNet's object detection and Mask R-CNN's segmentation capabilities creates a comprehensive solution for cotton crop area Estimation, fostering advancements in agricultural practices and ultimately contributing to sustainable and efficient crop management.

**DECLARATION**

**Conflict of Interests**

This manuscript does not disclose any conflicts of interest. To maintain our commitment to academic integrity, we have rigorously ensured the originality of our work to prevent any inadvertent entanglement with issues related to academic misconduct.

**Acknowledgment**

The invaluable support and resources generously provided by the esteemed Saveetha Institute of Medical and Technical Sciences played an indispensable role in the triumphant culmination of this study. The authors extend their deepest appreciation for the immeasurable assistance rendered by this distinguished institution.

**Authors Contribution**

Data gathering, analysis, and text creation were all actively participated in by author NT. Author RN, on the other hand, made an efficient contribution to the idea of the research, carried out data validation, and offered valuable criticism throughout the paper review process.

**Funding**

Thankful for the following organizations for providing financial support that enabled us to complete the study.

1. Saveetha School of Engineering.
2. Saveetha Institute of Medical and Technical Sciences.
3. Saveetha University.
4. Vimkes Technologies Pvt.Ltd., Chennai.

**REFERENCES**

[Achanta, Radhakrishna, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Sabine Süsstrunk. 2024. “SLIC Superpixels Compared to State-of-the-Art Superpixel Methods.” Accessed March 11, 2024.](http://paperpile.com/b/37gESw/AxJB) <https://ieeexplore.ieee.org/abstract/document/6205760/>[.](http://paperpile.com/b/37gESw/AxJB)

[“A Convolutional Neural Network Approach for Counting and Geolocating Citrus-Trees in UAV Multispectral Imagery.” 2020. *ISPRS Journal of Photogrammetry and Remote Sensing: Official Publication of the International Society for Photogrammetry and Remote Sensing*  160 (February): 97–106.](http://paperpile.com/b/37gESw/JAwq)

[“Assimilation of Remote Sensing into Crop Growth Models: Current Status and Perspectives.” 2019. *Agricultural and Forest Meteorology* 276-277 (October): 107609.](http://paperpile.com/b/37gESw/QSLt)

[“Combining UAV-Based Plant Height from Crop Surface Models, Visible, and near Infrared Vegetation Indices for Biomass Monitoring in Barley.” 2015. *International Journal of Applied Earth Observation and Geoinformation* 39 (July): 79–87.](http://paperpile.com/b/37gESw/pXjE)

[“Fine-Grained Maize Tassel Trait Characterization with Multi-View Representations.” 2015. *Computers and Electronics in Agriculture* 118 (October): 143–58.](http://paperpile.com/b/37gESw/JmaP)

[Ha, Jin Gwan, Hyeonjoon Moon, Jin Tae Kwak, Syed Ibrahim Hassan, Minh Dang, O New Lee, and Han Yong Park. 2017. “Deep Convolutional Neural Network for Classifying Fusarium Wilt of Radish from Unmanned Aerial Vehicles.” *Journal of Applied Remote Sensing* 11 (4): 042621.](http://paperpile.com/b/37gESw/L8pB)

[Ioffe, Sergey, and Christian Szegedy. 2015. “Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift.” In *International Conference on Machine Learning*, 448–56. PMLR.](http://paperpile.com/b/37gESw/a3lr)

[Kitano, Bruno T., Caio C. T. Mendes, André R. Geus, Henrique C. Oliveira, and Jefferson R. Souza. 2024. “Corn Plant Counting Using Deep Learning and UAV Images.” Accessed March 11, 2024.](http://paperpile.com/b/37gESw/w4cb)

<https://ieeexplore.ieee.org/abstract/document/8792201/>[.](http://paperpile.com/b/37gESw/w4cb)

[Korir, Nicholas Kibet, Jian Han, Lingfei Shangguan, Chen Wang, Emrul Kayesh, Yanyi Zhang, and Jinggui Fang. 2013. “Plant Variety and Cultivar Identification: Advances and Prospects.” *Critical Reviews in Biotechnology*, June. https://doi.org/](http://paperpile.com/b/37gESw/Nw91)[10.3109/07388551.2012.675314](http://dx.doi.org/10.3109/07388551.2012.675314)[.](http://paperpile.com/b/37gESw/Nw91)

[Lin, Tsung-Yi, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollar. 2017. “Focal Loss for Dense Object Detection.” In *Proceedings of the IEEE International Conference on Computer Vision*, 2980–88.](http://paperpile.com/b/37gESw/TKGp)

[Oh, Min-Hwan, Peder Olsen, and Karthikeyan Natesan Ramamurthy. 2019. “Counting and Segmenting Sorghum Heads.”](http://paperpile.com/b/37gESw/0vxj) <http://arxiv.org/abs/1905.13291>[.](http://paperpile.com/b/37gESw/0vxj)

[Santos, Anderson Aparecido dos, José Marcato Junior, Márcio Santos Araújo, David Robledo Di Martini, Everton Castelão Tetila, Henrique Lopes Siqueira, Camila Aoki, et al. 2019. “Assessment of CNN-Based Methods for Individual Tree Detection on Images Captured by RGB Cameras Attached to UAVs.” *Sensors*  19 (16): 3595.](http://paperpile.com/b/37gESw/my7J)

[Simonyan, Karen, and Andrew Zisserman. 2014. “Very Deep Convolutional Networks for Large-Scale Image Recognition.”](http://paperpile.com/b/37gESw/p6nT) <http://arxiv.org/abs/1409.1556>[.](http://paperpile.com/b/37gESw/p6nT)

[Szegedy, Christian, Sergey Ioffe, Vincent Vanhoucke, and Alexander Alemi. 2017. “Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning.” *Proceedings of the AAAI Conference on Artificial Intelligence* 31 (1). https://doi.org/](http://paperpile.com/b/37gESw/sBOX)[10.1609/aaai.v31i1.11231](http://dx.doi.org/10.1609/aaai.v31i1.11231)[.](http://paperpile.com/b/37gESw/sBOX)

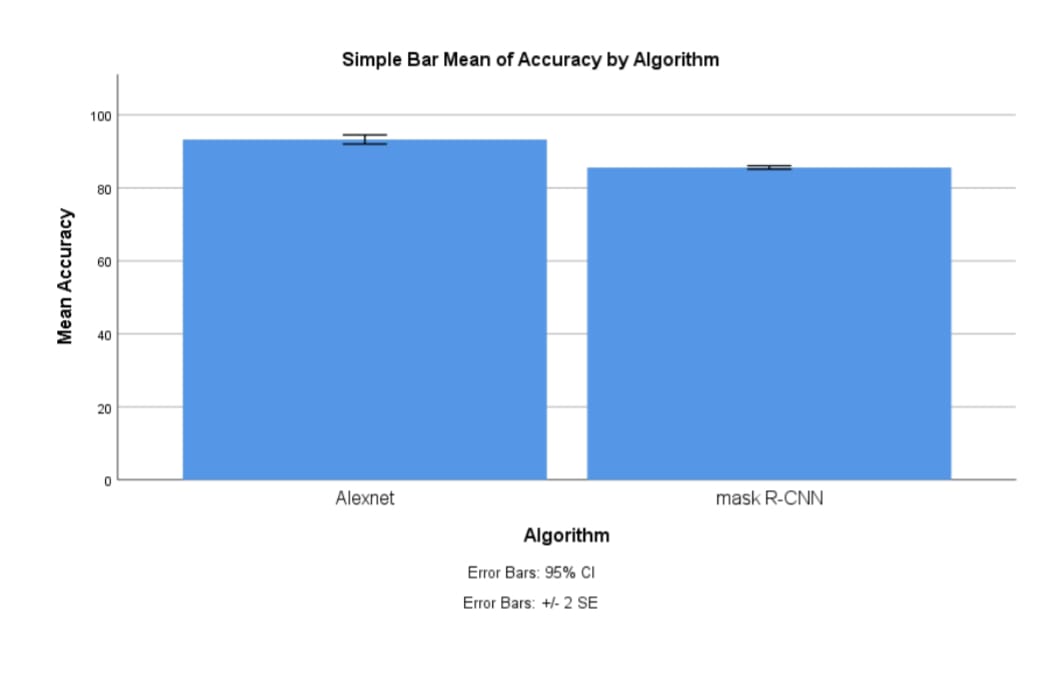
[Yu, Fisher, and Vladlen Koltun. 2015. “Multi-Scale Context Aggregation by Dilated Convolutions.”](http://paperpile.com/b/37gESw/kxW5) <http://arxiv.org/abs/1511.07122>[.](http://paperpile.com/b/37gESw/kxW5)

**Table 1:**Shows Statistical Analysis values of Mean accuracy (93.27), Standard Deviation(2.434), and Standard error deviation. The Alex net Algorithm and the Mask R-CNN algorithm have the values of the Mean accuracy, Standard Deviation, and Standard Error.

|  | **Algorithm** | **N** | **MEAN** | **STD.DEVIATION** | **STD.ERROR MEAN** |
| --- | --- | --- | --- | --- | --- |
| **Accuracy** | **Alexnet** | 15 | 93.27 | 2.434 | .628 |
| **Mask R-CNN** | 15 | 85.60 | .910 | .235 |

**Table 2.** Shows Comparison of Significance Level with value p<0.05. Both Alex net Algorithm and the Mask R-CNNAlgorithm have a confidence interval of 95% with the significance value 0.000( p<0.05).

| **Accuracy** | **Levene's test for equality of variances** | | **T-test for equality of means** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **F** | **sig** | **t** | **df** | **sig(2-tailed)** | **Mean difference** | **Std error difference** |  | |
| **Lower** | **Upper** |
| **LogisticRegression** | 20.671 | .000 | 11.427 | 28 | .000 | 7.667 | .671 | 6.292 | 9.041 |
| **Decision tree Algorithm** |  |  | 11.427 | 17.841 | .000 | 7.667 | .671 | 6.29 | 9.077 |



**Fig. 1**:Comparison of the Alex net Algorithm accuracy of (93.27%) and it has the mean accuracy of Mask R-CNNalgorithm (83.27%) . The mean accuracy of the Alex net Algorithm has no significant difference with the Random Mask R-CNN Algorithm with the significance value is 0.000 (p<0.05) . X Axis: Alex net Algorithm vs Mask R-CNN algorithm Y Axis: Mean accuracy ± 2 SD.