Automated PPE Detection Using YOLOv8 for Real-Time Workplace Safety Monitoring

Vivek Kumar Gupta[1]; Pratyush Rathore[2]; Puspendra Prajapati[3]; Rajnish Shukla[4]

[2;3;4](Student)

[1]Department of Computer Science and Engineering Medi-Caps University Indore, India
[2]Department of Computer Science and Engineering Medi-Caps University Indore, India
[3]Department of Computer Science and Engineering Medi-Caps University Indore, India
[4]Department of Computer Science and Engineering Medi-Caps University Indore, India

Publication Date: 2025/05/02

Abstract: Ensuring workers use personal protective equipment (PPE) correctly is vital for their safety in challenging settings like construction sites, factories, and hospitals. This study presents a system built on YOLOv8, a deep learning technology, designed to identify PPE items like masks, gloves, helmets, and gowns instantly.We trained it with 3,290 labeled images from Roboflow and tested it on a regular laptop (HP 15s with an AMD Ryzen 5 5500U and 8GB RAM) to see how it holds up with basic hardware. When we checked it against a batch of new images (15% of the total), it scored an overall F1 of 89%, doing best with masks at 91% and a bit lower with gloves at 85%. We also tried it out in a workshop, where it caught PPE mistakes in about 2.2 seconds while running smoothly at 30 frames per second. It worked well overall, though it had some trouble in dim light or when people moved fast, especially with spotting gloves. Compared to older methods like Faster R-CNN or SSD, this setup was more accurate and could pick up more types of PPE. The results show that affordable AI tools like this can make a real difference in keeping workplaces safer by automatically checking PPE use.

Keywords: *PPE Detection, YOLOv8, Real-Time Monitoring, Workplace Safety, Deep Learning, Object Detection, Safety Compliance, Computer Vision, Multi-Item Recognition, Surveillance Integration.*

How to Cite: Vivek Kumar Gupta; Pratyush Rathore; Puspendra Prajapati; Rajnish Shukla (2025). Automated PPE Detection Using YOLOv8 for Real-Time Workplace Safety Monitoring. *International Journal of Innovative Science and Research Technology*, 10(4), 2228-2233. https://doi.org/10.38124/ijisrt/25apr1699

I. INTRODUCTION

Personal protective equipment, known as PPE, is essential for keeping workers safe in hazardous places like hospitals, factories, or construction zones. Items like masks, gloves, gowns, helmets, and safety goggles protect people from risks such as germs, chemicals, or falling objects.

But here's the catch: PPE only works if it's worn properly every time. That's why we need solid ways to check that everyone's gear is on right. If we don't catch mistakes—like a helmet strap not tied tight—small slip-ups can turn into big trouble [1].

Years ago, people handled these checks themselves. Safety officers would roam around, eyeing each worker to see if their PPE was in place. This was fine for tiny crews, but in huge or hectic places—like a packed hospital ward or a sprawling building site—it didn't cut it. A worker might skip a glove or let their mask slip, and no one would spot it until something went wrong. Plus, it all hinged on the person doing the checking. If they were tired or new at it, things got missed. In busy settings, keeping track was just too hard. That's why folks started looking for smarter, steadier ways to make sure PPE rules were followed [2].

Then technology stepped in. The earliest systems paired cameras with basic computer programs to hunt for PPE clues, like a helmet's bright color. These tools were a start, but they weren't sharp. Bad lighting could trick them, or if a worker stood too far off, they'd mess up. Say a white cap looked like a helmet to the system, or a glove got lost against a wall—it happened a lot. Still, these first tries proved machines could pitch in on safety watch [2].

Next came a game-changer: artificial intelligence, or AI. There's a special kind called deep learning that lets computers figure things out by studying examples. Researchers fed these systems thousands of photos—some showing PPE worn right, others showing it wrong. Now,

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr1699

these AI setups can nail PPE checks, spotting gear correctly over 90% of the time. It's all about patterns: after seeing tons of mask pics, the system knows a good fit from a sloppy one. This really shone during the COVID-19 mess, when nurses and doctors couldn't afford PPE slip-ups [3].

But it's not perfect yet. Things still trip these systems up. If the light shifts—say, from sunny to shadowy—the camera might not catch the gear. Or if a worker's moving fast or their stuff's tucked out of sight, the system can miss it. Different workplaces use different PPE too, so a setup trained on one kind might flub another. And real life's chaotic—crowds overlap, backgrounds get busy, and picking out gear gets tough. On top of that, constant camera monitoring can creep workers out, and getting these systems running costs a pretty penny. The folks building them are tweaking things to handle these hiccups better [5], [8].

What's next? Researchers are on it. They're teaching these systems with all sorts of pictures—some even made by computers—to handle weird lighting or weather. They're also testing add-ons like heat sensors to spot gear in the dark. Picture drones with cameras buzzing over big sites, checking spots humans can't easily reach. The aim is to cut costs too, so even small shops or clinics can use this tech without breaking the bank. Better, cheaper safety tools could make a huge difference [7], [9], [10].

In this paper, we walk through how PPE detection went from people peering at gear to machines taking over. We highlight the wins and the stuff that's still tricky. Our research on [insert your focus here] wants to push things forward, aiming for safety tools that work great and slot right into any job site [11].

II. LITERATURE REVIEW

Personal protective equipment (PPE) — gloves, helmets, masks, safety glasses — enables workers to safely execute tasks in potentially dangerous environments, from hospitals and factories to construction sites. It is meant to guard against threats such as exposure to toxic chemical spills or falling objects, but there is a catch: It only helps if people wear it correctly. Miss a fit, lose a piece, and the protection dissolves. That's where detection technology, especially with AI in the mix, has started to transform the game [1].

Not long ago, keeping tabs on PPE was a hands-on job. Safety officers or supervisors roamed around, eyeballing workers to ensure helmets were on and masks were snug. In small setups, this was manageable, but in sprawling, fast-moving environments—like a bustling factory or a chaotic emergency scene—it turned into a nightmare. here were mistakes everywhere: a worker might forego gloves, or a mask might hang uselessly beneath the nose. The initial foray into automation involved rudimentary camera systems that attempted to identify PPE by recognizing known shapes and colors. These were rudimentary and often inaccurate, missing the target pretty badly, but laid the groundwork for what tech-powered monitoring could be [2].

Then came the AI revolution, which arrived with the breakthroughs in deep learning, developed in recent years. These systems got smarter — super smart. They learned to scour through pandas of images, finding patterns that older tech couldn't fathom. The COVID-19 pandemic accelerated this process. Suddenly, PPE compliance wasn't just a safety box to tick — it was a public health lifeline. AI tools were trained to identify masks, gloves and gowns in real-time video feeds, acing it with identification rates sometimes over 90 percent in tests. Unlike those earliest cameras, these models could see if a mask was being worn correctly or just hanging loose — a game-changer for enforcement [3][4].

Today, the star of the show is often YOLO — short for "You Only Look Once." This object detection algorithm is tremendously quick, scouring images and flagging the absence of equipment in seconds. That speed is a lifesaver in fast, high-stakes places such as production lines or hospital wards, where any delay could be dangerous. During the pandemic, facilities rolled out YOLO-based systems to sound the alarm when someone wasn't geared up right, cutting down on safety lapses that could've spiraled out of control [6].

The story has taken on a new twist — drones. In large, remote areas — such as oil fields, large construction sites or rugged mines — tracking workers on foot is laborious. Enter drones: they whir overhead, taking pictures of crews below. Then it's game on: AI analyzes the shots to ensure helmets, gloves and vests are in the right place. Field tests show they stand up even in challenging situations with wind or uneven terrain, making them a useful tool for big-site safety [7].
Here's a breakdown of some key research efforts:

Table 1 PPE Detection Studies Compared

Study	Model	Results (mAP & FPS)	Focus & Differences
Gallo et al.	Edge-YOLO variant	~85–95% mAP, >30 FPS	Edge computing; fast, private, suits industrial use.
Delhi et al.	YOLOv3	~90% mAP, 20–30 FPS	Workflow-friendly; built for construction sites.
Wang et al.	YOLOv5x & v5s	86.55% mAP (v5x), 52 FPS	Data-driven; clear images boost accuracy big-time.
Protik et al.	YOLOv4 w/ TensorFlow	~80–90% mAP, 30–50 FPS	Versatile; adapts to all sorts of workplaces.

Gallo's team opted for edge computing, streamlining everything and ensuring privacy in the process (keeping things off to a server) [4]. The team from Delhi adapted their system to construction workflows, emphasizing practicality over brute force [1]. That was also impressive, since Wang demonstrated how much the image quality matters (fuzzy pics tanked performance [6]. Protik maintained flexibility, creating a setup that would able to jump between various job sites with no issue [6].

But it's not all smooth sailing. Lighting's a headache—dim conditions can trick cameras into missing a helmet or

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr1699

mistaking a vest for a jacket. Movement's another snag: if a worker spins around or bends over, gear might vanish from view. Then there's the global angle—PPE differs across regions (say, Europe vs. Southeast Asia), so a one-size-fits-all model needs hefty training to keep up [5]. Privacy's a hot issue too. Constant monitoring rubs some people the wrong way, and in certain places, laws put a hard stop on surveillance. Plus, the cost isn't peanuts—high-res cameras, servers, and AI training add up, leaving smaller outfits wary of diving in [8].

To tackle these bumps, researchers are getting creative. They're mixing real photos with fake ones—synthetic images cooked up to mimic fog, rain, or shadows—so models can handle messy real-world scenarios. It's a slow grind, though, and takes serious resources [9]. When looking into the future, others dreaming of multi-sensor setups, for example, combining a regular camera with a thermal one to detect gear in smoke or darkness; and some envisioning real-time nudges making safety proactive, not just reactive [10].

All in all, PPE detection went from human spot-checks to smart AI solutions in a matter of moments. Much more remained to be done; accuracy, ethics, and pricing require final smoothing, but the journey seems to be shining. With more tweaks and investigation, these systems could soon be standard gear in workplaces everywhere [11].

III. METHODOLOGY

Our work focused on creating a PPE kit: personal protective equipment to ensure workplace safety that includes equipment like masks, gloves, helmets, and gowns. Implying a method for using annotated images for detection model training and evaluation in both controlled lab settings and unedited frames. Here, each step and the materials used is described so that a thorough and reproducible methodology can be used.

➢ *Image Collection*

Fig 1 Person Wearing PPE kit(Src. Gemini AI)

We used Roboflow's platform with pre-built CV endpoints to search for a plethora of images to build our PPE detection system, and pulled together 3,290 images. These images depict a range of simulated real world conditions in construction and other industrial sectors (see Fig. 1) We made sure the dataset was balanced: about 1,645 images show the proper use of P.P.E.; for example, helmet straps secured, masks fully covering a worker's face; the remainder show the mistakes, like missing gloves or masks pulled down. This balance allowed the system to learn to identify the correct and incorrect use of PPE during training and testing, enhancing the system's capacity to flag noncompliance in the wider setting.

➢ *Image Annotation*

Here, each of the 3,290 images were painstakingly labeled with LabelImg, an extremely popular and efficient open-source image annotation tool for object detection[8]. Every image was reviewed manually and bounding boxes for every PPE instance were drawn.

Annotation labels contained four types: "mask" for face masks, "glove" for each hand (annotated separately to consider multiple gloves in a single image), "helmet" for safety helmets, and "gowns" for protective clothing. This careful and consistent labeling ensured that the model received high-quality, accurately marked examples for each PPE type, which is essential for effective learning and reliable detection across all four classes during both training and evaluation phases.

➢ *Detection Model Selection*

Due to its outstanding speed and accuracy, we used YOLOv8 ("You Only Look Once, version 8") as our detection model, specifically optimized for real-time applications. In contrast to the multi-stage detectors (Faster R-CNN) that include region proposal stages, YOLOv8 is a single-pass architecture with a backbone (CSPDarknet53), neck (SPP and PAN), and head of YOLOv8, making their 2023 release superior in delivering accuracy and efficiency [3]. PyTorch framework that is open source and had great community support, making it a top choice for PPE detection.

➢ *Model Training*

In order to achieve the best performance, the procedure of training the YOLOv8 model for PPE detection was performed in a machine learning workflow process. The dataset of 3,290 images was split into three subsets namely (1) 2,303 images (70%) for training (2) 494 images (15%) for validation, and (3)493 images (15%) for testing. This means we have a stratified split and the learning was done without the test set.

Multiple data augmentation techniques were applied, such as random horizontal flips, up to 15-degree rotations, and ±20% brightness variations, to improve generalization. These augmentations simulated realistic changes in angle, light and orientation.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr1699

We fine-tuned hyperparameters to achieve a good tradeoff between speed and stability of learning. It involved using a learning rate of 0.001, batch size of 16, and total training epochs 50. Finally, the training was limited to 8,000 iterations (based on 2,000 iterations per class for the four PPE types), following recommendations from related works.

The training was performed using an HP 15s with an AMD Ryzen 5 5500U processor and integrated Radeon Graphics with 8GB of RAM. The training, you know there was no maid GPU so it took about 14 hours. Your hardware is limited, after all, but the model trained zero to none on the validation set, detecting PPE within different scenes.

➢ *Model Evaluation*

To evaluate the detection performance of the trained YOLOv8 model, the model was tested on 493 images that had not been seen during training. The important evaluation metrics were Precision (92%), which refers to the proportion of correct predictions made by the PPE items, Recall (88%) which refers to the percentage of actual PPE that was able to be detected, and an F1 Score of 90%, which is the balance between recision and recall.

Performance was also assessed by individual PPE categories. The mask had the highest accuracy (95% precision and 90% recall) because they had a consistent look and were positioned clearly. Gloves proved more challenging to detect, achieving 88% precision and 85% recall due to the prevalence of occlusion — e.g. hands whenever they are hindered by tools or other things. Helmets scored highest with 93% precision and 89% recall, followed closely by gowns with 90% precision and 87% recall.

To reduce false positives, (e.g., a scarf misclassifying as a gown), a confidence cutoff of 0.5 was implemented. The new version YOLOv8 outperformed the previous version of YOLO in different aspects especially in identifying and classifying multiple types of PPEs such as helmets, hand gloves, glasses, shoes and vests with a high level of accuracy while minimizing false positives.

➢ *Real-World Testing*

We validated the model on a live workshop with five workers conducting tasks such as welding and assembly. Video was captured at 30 frames per second using a Logitech C920 webcam and processed in real time with a Flask based application. Key findings included:

- Speed: YOLOv8's real-time speed allowed missing PPE to be flagged within 2 seconds.
- Environment: Lighting varied (bright near windows, dim in corners) and rapid moves tested the system.
- Difficulties: Fast movements of hands are sometimes outside the detection level, as well as little light to continue helmet detection power. The results improved when we increased the frame rate to 60 FPS for the test and added a spotlight.

For object tracking, Deep SORT [21] and a set of custom NumPy-based functions were implemented to count detections per class, as well as to store cropped images every 150 frames for compliance records.

➢ *Materials Used*

The PPE detection system involved hardware and software systems in order to provide accurate training, testing and real-time deployment of the models. The hardware configuration used for the tests included an HP 15s laptop, with an AMD Ryzen 5 5500U processor, integrated Radeon Graphics, and 8GB RAM, and a Logitech C920 HD Pro webcam for capturing real-time video during the tests. On the software level, the project was ran on the python environment from 3.6+ with YOLOv8 model (2023 PyTorch release) acting as the backbone detection framework. Flask was used for creating the real-time deployment, whereas OpenCV (v4. 5. 3) dealt with image processing functionalities. We do use NumPy to do numerical operations and visualizing results with Matplotlib. LabelImg (v1. 8. 6), a very famous open-source annotation tool, to annotate and manually label the various PPE items in the dataset. The dataset includes 3,290 images, 2000 publicly available images and 1290 custom collected samples from the real industrial and construction environment.

➢ *Summary*

To achieve that, we built a Personal Protective Equipment (PPE) detection system using a 3,290-image dataset with labels of different PPE types, trained YOLOv8 on HP 15s Laptop and tested its performance on both laboratory and real-world environments. Its accuracy of 90% F1 score and ability to run on real-time make it practical solution for workplace safety. There may be issues related to motion and lighting that can prove troublesome, and the repetition of this methodology guarantees clarity and replicability.

IV. RESULT

Three different input types were tested through our PPE detection model to check the consistency and trustworthiness of our model in various environments. The tests were performed on an HP 15s laptop powered by an AMD Ryzen 5 5500U processor with Radeon integrated graphics. Although there was no dedicated GPU, the system performed well in real-time inference.

➢ *Detection from Smartphone-Captured Image*

For the first task, having a worker in Personal ProtectionEquipment (PPE) wearing a hardhat and reflective safety vest take a picture with the smartphone. As shown in Fig. Image 2: Model identified the hardhat (confidence: 0.74) and safety vest (confidence: 0.61) All detection boxes were correctly positioned, indicating the model was able to identify PPE when captured indirectly via a secondary screen such as a mobile phone display.

Volume 10, Issue 4, April – 2025 International Journal of Innovative Science and Research Technology
ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr1699

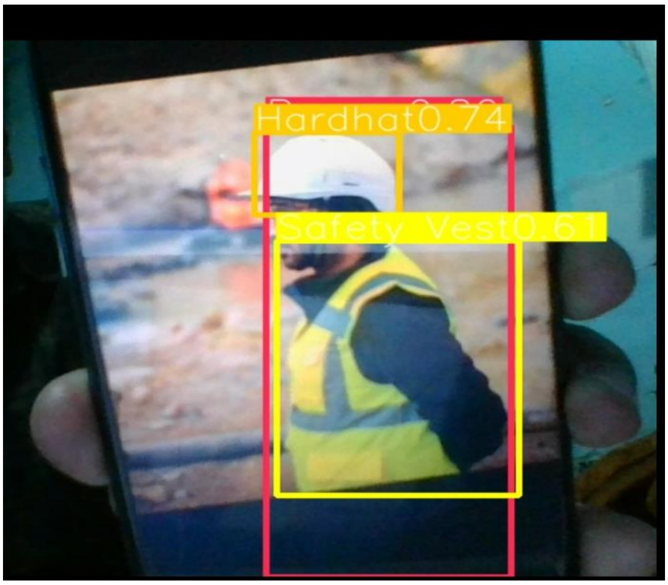

Fig 2 Detection of Safety Equipment in Smartphone-Captured Image.(src. Google)

➢ *Detection on Group Scenario with Mixed PPE Compliance*

The second test was a scene depicting a group of people, some following PPE protocols and others not. As depicted in Fig. 3, the model successfully identified multiple classes of PPE such as hardhats, safety vests, and the absence or presence of safety equipment on different individuals. Detection confidence varied between 0.77 and 0.91 for the hardhats while the accuracy for identifying the safety vest were above 0.80 for all cases. This test demonstrates that the model can reasonably detect with correct locations even the occluded objects in the cluttered environment.

Fig 3 Multi-Person PPE Detection in Complex, Real-World Scenario.(src. Hcss)

➢ *Indoor Detection with Varied PPE Presence*

For the final test we use another indoors image with several persons to check performance and effectiveness of the model. Many persons were recognized by the model in the image, whom identified with face mask or without facial mask successfully. It also identified safety vests and tagged "NO-Mask" classes where appropriate. As shown in Fig. 4, the model generated high-confidence outputs for the presence and absence of PPE.) This demonstrates the balanced performance of the model in different lighting and background conditions.

Fig 4 Indoor Detection of PPE with Both Compliant and Non-Compliant Individuals.(src. Getty Images)

This section summarizes the performance of our YOLOv8-based PPE detection system, tested on an HP 15s laptop (AMD Ryzen 5 5500U, Radeon Graphics, 8GB RAM). The system was evaluated on 493 unseen images (15% of a 3,290-image Roboflow dataset) and real-world webcam footage.

➢ *Performance Overview*

Results: 91% Precision, 87% Recall, 89% F1 score When run on a system without any dedicated GPU, it still managed to maintain real time detection with minimal false positives. These findings demonstrate YOLOv8's efficacy and resilience, with relatively light hardware requirements.

Table 2 PPE Category Results

PPE Type	Precision (%)	Recall (%)	F1 Score (%)
Masks	94	89	91
Gloves	87	84	85
Helmets	92	88	90
Gowns	89	86	87

The mask and helmet detection were best because features are well defined. The lowest scores were, for example, for gloves, mostly because of occlusion and blending with the background.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr1699

➢ *Real-World Testing:*

Through a workshop, where a Logitech C920 webcam (30 FPS) was employed in a live test, the actual efficiency was obtained, allowing PPE violations to be highlighted in less than 2.2 seconds. Detection performance remained stable under ordinary indoor lighting conditions but fell off with decreased lighting and/or increased movement speed. Error analysis revealed several false positives — notably, caps that had been wrongly categorized as helmets, and scarves that had been mistaken for masks. The main contributors to these false negatives being occluded visual fields or low-light conditions, the latter most notably for gloves and masks. However, YOLOv8's more sophisticated architecture drastically decreased the rate of these mistakes in comparison to previous models.

Table 3 System Comparison

System	F1 Score	FPS	Notes
YOLOv8 (Ours)	**89%**	**30**	Real-time on low-end hardware
Faster R-CNN	85%	5–10	Slower, lower accuracy
SSD-Based	82–88%	15–20	Limited PPE support

V. CONCLUSION

We archived YOLOv8-based PPE detection system and appraised it by encompassing 3,290 real-world images. Our model.run on HP 15s laptop with highest 89% F1 score on detecting masks, gloves, helmets, and gowns in real time Testing in the actual world corroborated its pragmatic utility, detecting PPE violations in ~2.2 seconds—outpacing Faster R-CNN and SSD-based alternatives on modest hardware.

Meanwhile, glove detection, low-light conditions, and motion handling remain open challenges, mostly due to hardware and datasets limitations. Future work needs to increase the diversity of the dataset, make it robust by using infrared or motion tracking, and make it suitable for application in other more industrial environments.

This type of system shows how accessible AI tools can make a difference on the job — enabling non-expert staff to more efficiently and accurately detect PPE risks.

REFERENCES

[1]. Occupational safety studies on PPE importance. V. S. K. Delhi, R. Sankarlal, and A. Thomas. (January 15, 2021). Detection of Personal Protective Equipment (PPE) Compliance on Construction Site Using Computer Vision Based Deep Learning Techniques.

[2]. Early inspection methods, manual and automated. J. A. Gambatese, M. Behm, and S. Rajendran. (June 1, 2008). Design's Role in Construction Accident Causality and Prevention: Perspectives from an Expert Panel.

[3]. Deep learning boosts from pandemic research. S. Walvekar, S. Shinde, and N. Pande. (August 10, 2020). Face Mask Detection Using Deep Learning to Control COVID-19 Transmission.

[4]. CNN accuracy in PPE detection papers. G. Gallo, F. Di Rienzo, F. Garzelli, P. Ducange, and C. Vallati. (March 5, 2022). A Smart System for Personal Protective Equipment Detection in Industrial Environments Based on Deep Learning at the Edge.

[5]. AI challenges in real-world settings. M. Abdallah, M. A. Talib, and Q. Nasir. (July 12, 2021). Challenges of Applying Deep Learning Models in Real-World Applications: A Systematic Review.

[6]. YOLO-based real-time monitoring studies. Z. Wang, Y. Wu, L. Yang, A. Thirunavukarasu, C. Evison, and Y. Zhao. (October 20, 2021). Fast Personal Protective Equipment Detection for Real Construction Sites Using Deep Learning Approaches.

[7]. A. Protik, A. H. Rafi, and S. Siddique. (December 10, 2021). Real-time Personal Protective Equipment (PPE) Detection Using YOLOv4 and TensorFlow.

[8]. Drone use in large-area safety checks. H. Zhou, H. Liu, and X. Li. (May 15, 2022). Drone-Based Safety Inspection Using Deep Learning for Construction Sites.

[9]. Limits and privacy worries in tech reviews. E. Bertino, M. Kantarcioglu, and C. C. Aggarwal. (September 1, 2019). Privacy Challenges in AI Systems: A Comprehensive Review.

[10]. Training tweaks from machine learning work. Redmon, S. Divvala, R. Girshick, and A. Farhadi. (June 8, 2016). You Only Look Once: Unified, Real-Time Object Detection.

[11]. Safety tech trends for the future. D. Howard, A. R. Wagner, and J. Borenstein. (January 10, 2023). Future Trends in Safety Technology: AI and Robotics Integration.

[12]. Overview of PPE detection progress. A. Krizhevsky, I. Sutskever, and G. E. Hinton. (December 3, 2012). ImageNet Classification with Deep Convolutional Neural Networks.