Technical White Paper

Advanced AI Vision Processing Using AM69A for Smart Camera Applications

Do-Kyoung Kwon, Karan Saxena, and Manisha Agrawal

ABSTRACT

The AM69A processor is the high-end edge AI device among the AM6xA scalable embedded processor family. This processor is designed for edge AI with as many as 12 cameras and provides various ways to optimize the performance at the network edge with heterogeneous processing cores and integrated hardware accelerators. The AM69A processor is the best embedded processor in the market for the industrial applications, where multiple camera inputs are processed for high-performance analytics at low power and low latency. The Processor SDK Linux for AM6xA (PSDK Linux) and tools provided make the development of the multi-camera AI applications simpler and faster on the AM6xA family including AM69A by taking advantage of vision and AI hardware accelerators. The PSDK Linux provides the unified software development framework for the AM6xA processor family so that the development experience on AM62A or AM68A makes the development process on AM69A much easier.

Table of Contents

List of Figures

List of Tables

Trademarks

Arm® and Cortex® are registered trademarks of Arm Limited.
Linux® is a registered trademark of Linus Torvalds.
Python® is a registered trademark of Python Software Foundation.
All trademarks are the property of their respective owners.

1 Introduction

A camera is a primary sensor modality for robots and machines to perceive and comprehend environments around them because of the rich visual data that cameras provide. The advance of deep learning-based AI and the embedded processors with AI capability, that is, edge AI processors, have made it possible to analyze enormous and complex visual data with higher accuracy but at lower power than before. Consequently, cameras are the most widely used sensors to analyze scenes, detect obstacles, recognize tags, and 2D and 3D barcodes, localize the position of objects as well as the position of the ego robot, map environment, and so forth.

Depending on the proximity of compute resources and data sources, there are two approaches to execute AI for video analytics, cloud AI and edge AI. Cloud AI processes visual data on the central computing infrastructure for training and inference of deep neural network (DNN) model. Cloud AI can analyze enormous amounts of data using substantial computing resource and therefore cloud AI has been dominating – especially in model training. However, since data needs to be transferred to the cloud, cloud AI presents the latency and security issues for real-time applications. In contrast, edge AI runs DNN model inference on the devices directly connected to camera sensors. Since camera data is processed locally, edge AI enables real-time processing with the reduced latency and security issues.

Edge AI requires the low-power edge AI processor that can process multiple cameras and execute multiple DNN inferences on them simultaneously. As edge AI processors become more powerful, edge AI technology is getting widely used in many applications, which in turn poses challenges to the edge AI processor in terms of size, power consumption and heat dissipation. The processor needs to fit in a small form factor and operate efficiently under the harsh environment of factories and construction sites as well as inside vehicles or cameras installed on roads. Moreover, certain equipment such as mobile machines and robots necessitates the edge AI processor to be certified for the applications that follow strict functional safety standards.

This paper introduces the highly-integrated AM69A processor. Several use cases of edge AI running on the AM69A are presented with the estimations of resource utilization and power consumption. The use cases include AI-Box, machine vision, multi-camera AI, and so on. Developing the edge AI systems using the heterogeneous architecture of the AM69A with the optimized AI models and the easy-to-use software architecture is also discussed.

2 AM69 Processor

The AM69A processor is the best performance device among the AM6xA scalable embedded processor family. Along with the octal-core Arm® Cortex® A72 microprocessor, the AM69A provides the most significant levels of processing power, image and video processing, and graphics capability. Compared with the AM62A(1) and the AM68A(2), which are excellent choices for the applications with 1 – 2 cameras and 4 – 8 cameras, respectively, the AM69A enables the real-time processing of 12 cameras with improved AI performance. As shown in Figure 2-1, the AM69A processor features the following multiple sub-systems based on the heterogeneous architecture:

- **An octal-core Arm Cortex-A72 microprocessor** at 2 GHz provides up to 100K Dhrystone Million Instructions Per Second (DMIPS).
- **Vision Processing Accelerator V3 (VPAC3)** performs image processing in Vision Image Sub-System (VISS) to support raw image sensor through de-mosaic, defective pixel correction, auto exposure, auto white balance, chromatic aberration correction (CAC), and so forth. In addition, VPAC3 includes Lens Distortion Correction (LDC), Multi-Scaler (MSC), and Bilateral Noise Filter (BNF) hardware accelerators (HWAs) to accelerate correction of distorted images, down scaling of images into multiple resolutions and noise filtering, respectively. The AM69A has two instances of VPAC3, which can process 1,200 MP per second (MP/s) when assuming 20% system overhead.
- **Digital Signal Processing (DSP) and Matrix Multiplication Accelerator (MMA)** are integrated together for DL acceleration as well as traditional computer vision tasks. The AM69A processor has four 512-bit C7x DSP running at 1 GHz. And each of them is tightly coupled with one of four MMAs capable of 4K (64 × 64) 8-bit fixed multiply accumulates per cycles. When run at 1 GHz, four MMAs provide a maximum compute speed of 32 Trillion Operations per Second (TOPS).

- **H.264, H.265 codec** can encode and decode multiple channels simultaneously. H.264, H.265 codec supports H.264 Baseline, Main, High Profile at L5.2, and H.265 Main Profile at L5.1. There are two instances of video codec so that the H.264, H.265 encoder and decoder can process 960 MP/s, for example, 16 channels of 2MP at 30 frames per second (fps).
- **3x 4-lane mobile industry processor interface (MIPI) CIS-2 RX** ports are equipped in AM69A. Three high-resolution (for example, 12MP) cameras can be directly connected to CSI-2 RX ports, captured and pre-processed by two VPAC3 instances. Capturing twelve 2MP cameras is possible via MIPI CSI-2 4-to-1 aggregators.
- **BXS-4-64 GPU** offers up to 50 Giga Floating-point Operations Per Second (GFLOPS) to enable dynamic 2D and 3D rendering for enhanced viewing applications.
- **Display Sub-System (DSS)** supports multiple displays with the flexibility to interface with different panel types such as eDP, DSI, and DPI.
- **Improved memory architecture and high-speed interfaces** improve the system throughput and energy efficiency by enabling high utilization of cores and HWAs. AM69A supports up to 64 Giga Bytes Per Second (GBps) DDR memory bandwidth.

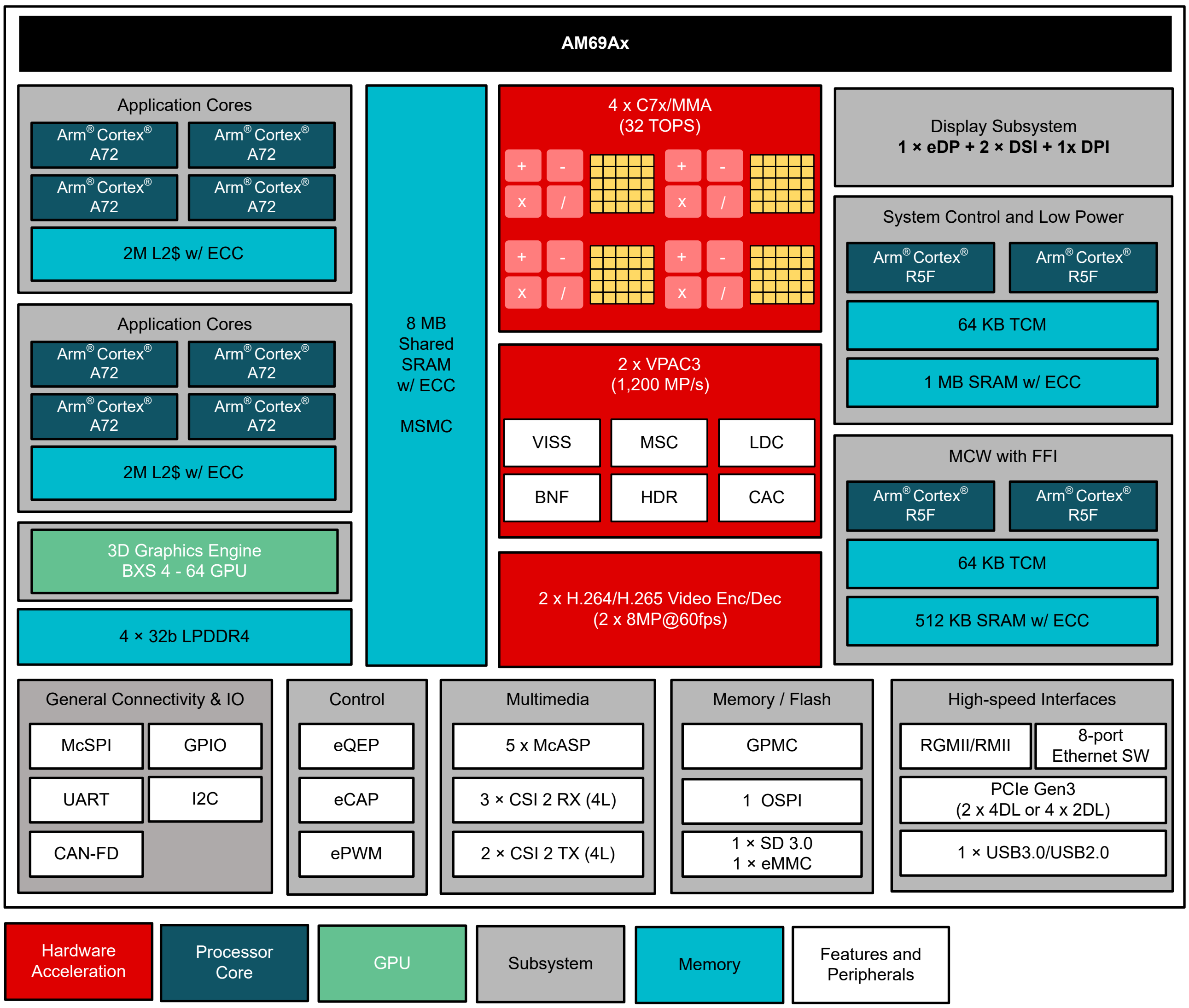

Figure 2-1. AM69A Block Diagram With Subsystems

Deep learning inference efficiency is crucial for the performance of an edge AI system. As the *Performance and efficiency benchmarking with TDA4 Edge AI processors* application note shows, MMA-based deep learning inference is 60% more efficient than a GPU-based one in terms of FPS and TOPS. The optimized network models for C7xMMA are also provided by TI Model Zoo[(3)], which is a large collection of DNN models optimized for C7xMMA for various computer vision tasks. The models include popular image classification, 2D and 3D object detection, semantic segmentation, and 6D pose estimation models. For the several models in TI Model Zoo, the 8-bit fixed-point inference performances on the TI embedded processors including AM69A can be evaluated via TI's Edge AI Studio.

The multicore heterogeneous architecture of AM6xA provides flexibility to optimize the performance of an edge AI system for various applications by utilizing suitable programmable cores or HWAs for particular tasks. For example, on AM69A, computationally-intense deep learning (DL) inference can run on four instances of MMA with optimized DL models, and vision processing, video encoding and decoding can be offloaded to two instances of VPAC3 and hardware-accelerated video codec for the best performance. Other functional blocks can be programmed in eight A72 cores or available C7x cores. Section 3 describes how edge AI systems can be built on AM69A for several industrial use cases.

3 Edge AI Use Cases on AM69A

The popularity of edge AI technology is increasing in many existing and new use cases. The AM6xA scalable processor family is well designed for them thanks to the multicore heterogeneous architecture. This section introduces popular use cases of edge AI with different input requirements, for example, resolution, frame rate, and task and computation requirements. The distribution of tasks among multiple cores and HWAs in the AM69A is described to optimize the performance. The use cases introduced in this section run on AM62A and AM68A as well, but with a reduced number of bitstreams and camera inputs.

3.1 AI Box

AI Box is a cost-effective way of adding intelligence to existing non-analytics cameras in retail stores, traffic roads, factories and buildings. AI Box is preferred over AI camera because AI Box is more cost effective than replacing legacy cameras with smart cameras that have AI capabilities. Such a system receives live video streams from multiple cameras, decodes them and does intelligent video analytics at the edge relieving the burden of transferring large video streams back to the cloud for analysis. The video analytics outputs are encoded before being streamed out and saved in storage. The exemplary applications of AI Box include

- **Security surveillance system**, which detects events and anomalous activities in the areas monitored by the remote cameras.
- **Smart traffic management system**, where AI Box is running deep learning networks for vehicle counting, vehicle type classification and moving direction predictions for traffic flow measurement.
- **Workplace safety system**, which monitors workplace to provide compliance with all safety standards imposed, for example, workers wearing personal protective equipment (PPE).

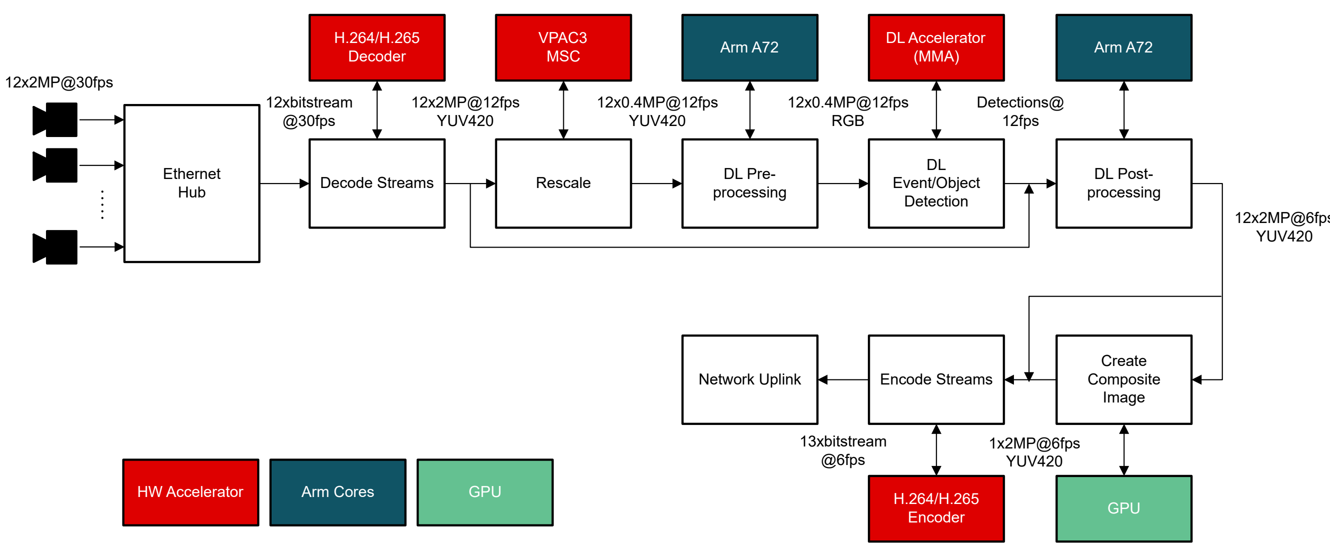

Figure 3-1. AI Box Block Diagram With Data Flow on AM69A

Figure 3-1 shows the data flow for AI Box on AM69A, where 12 channels of 2MP bitstreams are coming through Ethernet at 30 fps. The hardware accelerated H.264 or H.265 decoder decodes the bitstreams and the decoded frames are scaled to smaller resolution by MSC. DL networks run on these smaller-resolution frames at lower frame rate, for example, 12 fps. In DL pre-processing, the smaller-resolution frames in YUV are converted to RGB, which is the input format to DL network. MMA accelerates DL inference. In DL post-processing, the network outputs are overlaid on the input frame. Then the output frames from 12 channels are stitched together into a single 2MP frame and 13 channels, that is, 12 channels plus 1 composite channel, are encoded by hardware accelerated H.264 or H.265 encoder at lower frame rate, for example, 6 fps, and streamed out or saved in storage.

Table 3-1 summarizes the resource utilization and estimated power consumptions with 12 and 16 channels of bitstreams. However, due to the limited maximum throughput of the video codec, the input frame rate and output frame rate need to be reduced to 24 fps and 4 fps, respectively, for 16 channels of bitstreams. An assumption made here is that each channel needs 1 TOPS for inference. Two C7x cores are still available for additional vision processing and JPEG image encode to create snapshots. While both DL pre- and post-processing run on A72 cores in this example, they also can run on the available C7x cores, in which power consumption can be slightly different.

Table 3-1. AM69A Resource Utilization and Power Consumption Estimate for the AI Box Use Case

Main IP	Utilization (12 × 2MP at 30 fps)	Utilization (16 × 2MP at 24 fps)
Decoder	12 × 2MP at 30 fps = 720 MP/s (75%)	16 × 2MP at 24 fps = 768 MP/s (80%)
Encoder	12 × 2MP at 6 fps + 1 composite × 2MP at 6 fps = 156 MP/s (17%)	16 × 2MP at 4 fps + 1 composite × 2MP at 4 fps = 136 MP/s (15%)
Encoder + Decoder	720 MP/s + 156 MP/s = 876 MP/s (92%)	768 MP/s + 136 MP/s = 908 MP/s (95%)
GPU	20%	20%
VPAC (MSC)	12 × 2MP at 30 fps = 720 MP/s (60%)	16 × 2MP at 24 fps = 768 MP/s (64%)
MMA	12 × 1 TOPS per ch = 12 TOPS (38%)	16 × 1 TOPS per ch = 16 TOPS (50%)
8 × A72	DL pre- and post-processing, depacketization, JPEC encode, and so forth (50%)	DL pre- and post-processing, depacketization, JPEG encode, and so forth (40%)
DDR Bandwidth	9.49GBps (14%)	11.95GBps (18%)
Power Consumption (85°C)	18 W	18 W

3.2 Machine Vision

Industry 4.0 targets increased automation for production processes within the manufacturing industry. And Industry 5.0 emphasizes the human-centric collaboration between human and robots with artificial intelligence, that is, collaborative robot (cobot), to optimize the manufacturing process with improved automation. Machine vision is one of key technologies in Industry 4.0 and 5.0 and the real-time processing of visual data at the edge is crucial for machine vision. The main use cases of machine vision include:

- **Quantity and presence inspection** where 2D or 3D cameras verify the presence or absence of parts and ingredients in assembly and packaging systems. Vision-based deep learning is used to detect, classify, and count parts and ingredients.
- **Visual quality inspection** where 2D or 3D vision-based deep learning is used for various visual inspection purposes, for example, detecting defects or identifying the characters on printed circuit boards (PCB), gauging the dimension of parts, verifying proper assembly of parts and the wrapping of labels around containers, detecting tool wear defects as preventive maintenance, and UAV- or drone-based fault detection systems of solar panels, turbines, pipelines, and so forth.
- **Vision-guided robot**, for example, the robot arm picking and placing parts or bins is another use case of machine vision for the improved collaboration between human and cobots. Using the camera mounted on the robot arm, the pose of the object to pick is estimated and the optimum path is calculated for the robot arm.
- **Camera-based barcode reading system** is gaining popularity in the e-commerce logistics market. Since many customers make purchases online with 1- or 2-day shipping or less, the logistic market has been widely adopting the camera-based barcode reading system to improve the package processing throughput and reduce average shipping time as a result. In the barcode reading system, packages are placed on the barcode scanning tunnel that is moving fast and the camera-based barcode readers mounted on the upper, left, and right sides of the tunnel read the 2D or 3D barcodes on the packages at a high frame rate. The highly successful barcode reading rate is important in the barcode reading system since a failure to read barcodes results in manual handling. For the package whose barcodes were not read, the operator manually types the information and replaces the damaged barcode with a new barcode. The manual interruption increases labor costs and reduces the efficiency resulting in poor throughput. To improve successful barcode

read rate, the camera-based barcode readers make use of AI to locate barcode and filter damaged or poorly captured barcodes.

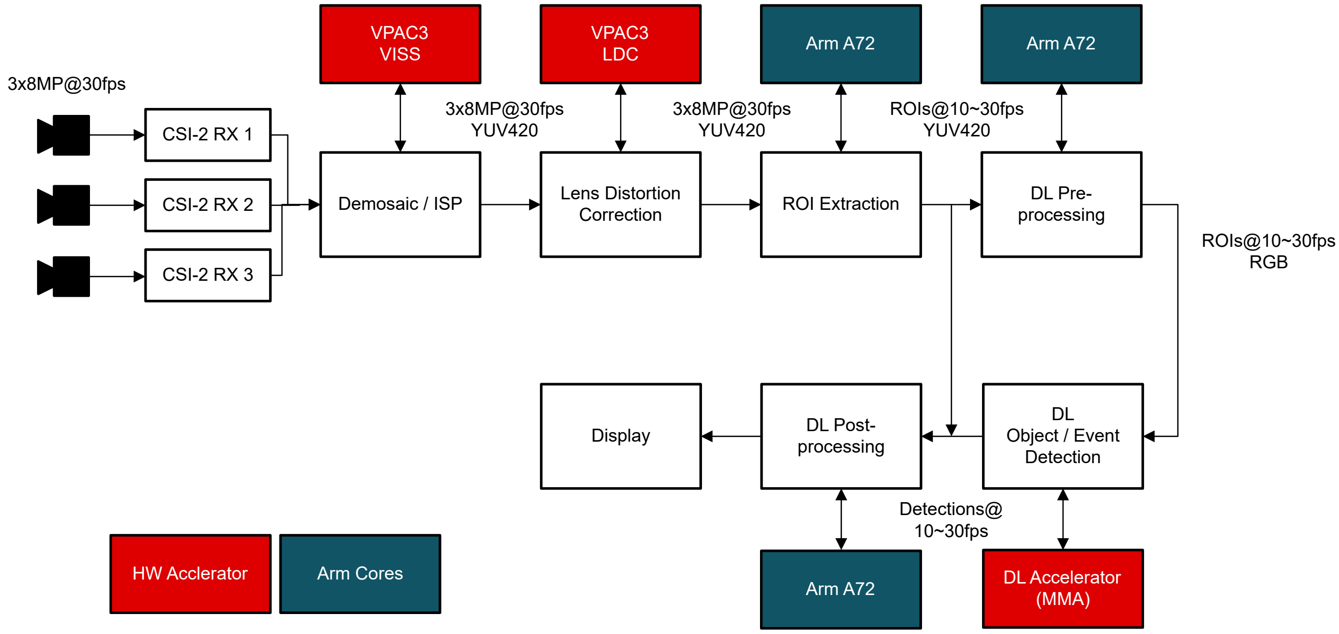

Figure 3-2. Machine Vision Block Diagram With Data Flow on AM69A

Figure 3-2 illustrates the data flow for a machine vision use case example on the AM69A, which involves capturing three 8MP image sequences at 30 fps through the MIPI CSI-2 RX ports. The captured raw Bayer images are processed and demosaiced to YUV by VPAC3 VISS, and VPAC3 LDC corrects any lens distortion that is present. In this machine vision use case, the DL networks are applied to region of interests (ROIs), which are extracted using A72 cores. The number of ROIs and their sizes vary depending on the specific use case. The frame rate at which DL networks are applied is also dependent on the use case. The outputs obtained through DL pre-processing, DL network on MMA, and DL post-processing are displayed via DSS. In the event of any unexpected detection, an alarm can be activated for human attention.

The resource utilization and estimated power consumption of AM69A are shown in Table 3-2 for three 8MP cameras. Single MMA is assumed to be fully utilized for one 8MP camera in this table even though the actual MMA utilization depends on the application. There is still enough room for CSI-2, VPAC, A72, and DDR bandwidth to process higher resolution. Therefore, the AM69A can enable the machine vision use case with higher resolution cameras, for example, 12MP, as long as MMA can handle the necessary DL inference at the cost of increased power.

Table 3-2. AM69A Resource Utilization and Power Consumption Estimate for the Machine Vision Use Case

Main IP	Utilization (3 × 8MP at 30 fps)
3 × CSI-2 RX	3 × 8MP at 30 fps = 11.52 Gbps (38%)
VPAC (VISS, LDC)	3 × 8MP at 30 fps = 720 MP/s (60%)
MMA	24 TOPS (75%)
8 × A72	ROI extraction, DL pre- and post-processing, and so forth (50%)
DSS	100%
DDR Bandwidth	15.35GBps (24%)
Power Consumption (85°C)	19 W

3.3 Multi-Camera AI

Lots of existing and emerging markets for edge AI processors fall into the multi-camera AI use case. Multi-camera AI is similar to AI Box. The difference is that multiple cameras are directly connected to the system through MIPI CSI-2 in multi-camera AI, while encoded bitstreams from remote cameras are streamed into the system through Ethernet in AI Box. The use cases of multi-camera AI include but are not limited to the following applications:

- **Surveillance camera** is one of the most popular multi-camera AI use cases. The applications of surveillance cameras include security, traffic monitoring, and so forth. For security, cameras monitor and record a specific area such as in a home, workplace, or public places to maintain safety of the area. For traffic monitoring, cameras are mounted at interactions, school zones, and frequently congested roads, and monitor traffic flow and accidents for the purpose of traffic management. In general, a surveillance camera system is connected to a network and the activities recorded are encoded and uploaded to the cloud for saving and for remote viewing.
- **Mobile DVR** is widely used to record the video footage from the cameras mounted inside and outside vehicles, which is useful for in-cabin monitoring, theft protection, and evidence in case of accidents. While mobile DVR is very similar to the surveillance camera system, mobile DVR saves the recorded video footage in the local storage devices such as a secure digital (SD) card or a solid-state drive (SSD).

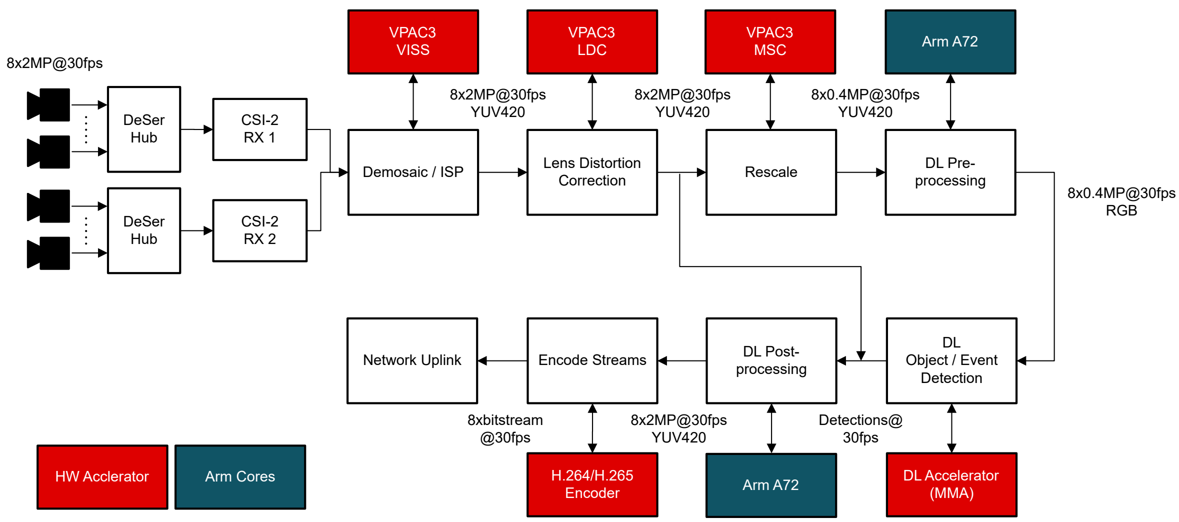

Figure 3-3. Multi-Camera AI Block Diagram With Data Flow on AM69A

Figure 3-3 shows the data flow for a multi-camera AI use case example on the AM69A. In this example, eight 2MP cameras are combined using two MIPI CSI-2 aggregators and captured at 30 fps through two MIPI CSI-2 RX ports, which is the main difference from AI Box. The images through ISP (Image Signal Processor), demosaic and lens distortion correction are scaled to smaller resolution by MSC. The outputs obtained through DL pre-processing, DL network on MMA, and DL post-processing are encoded by the hardware accelerated H.264 or H.265 encoder, and streamed out or saved in storage. Table 3-3 shows the resource utilization and estimated power consumption for the multi-camera AI use case with eight channels of a 2MP camera.

Table 3-3. AM69A Resource Utilization and Power Consumption Estimate for the Multi-Camera AI Use Case

Main IP	Utilization (8 × 2MP at 30 fps)
2 × CSI-2 RX	8 × 2MP at 30 fps = 7.68Gbps (26%)
VPAC (VISS, MSC, LDC)	8 × 2MP at 30 fps = 480 MP/s (40%)
MMA	24 TOPS (75%)
Encoder	8 × 2MP at 30 fps = 480 MP/s (55%)
8 × A72	DL pre- and post-processing, and so forth (50%)
DDR Bandwidth	15.13GBps (22%)
Power Consumption (85°C)	20 W

3.4 Other Use Cases

Other edge AI use cases, for which the AM69 processor is designed, include the followings:

- **Smart shopping cart** is an emerging end equipment for enhanced shopping experiences. Multiple cameras are mounted on the shopping cart to automatically detect the items placed in the cart, read barcodes on them, understand any activity in the cart afterward, calculate order totals, and allow consumers to pay for groceries, bypassing long checkout lines. In a barcode reader, deep learning networks are used to locate barcodes and filter damaged or poorly captured barcodes to improve the barcode read rate. Additional cameras are mounted for localization, for example, using AprilTags printed on the ground or ceiling. With the localization of a shopping cart, a more personalized shopping experience can be provided by identifying the locations of items on the shopping list and also recommending new products to customers.
- **Functional safe 3D perception** provides advanced assistance to machine operators for preventing collision and protecting nearby workers and pedestrians. Multiple cameras are mounted around vehicles and mobile machines such as the ones used for construction, agriculture, and mining. Next AI based 3D perception with functional safety is enabled. When combined with localization and navigation, the technology is enabling mobile machines and cobots to operate fully autonomously yet co-exist alongside humans and properties increasing the task efficiency.
- **Smart agriculture and farming** are also emerging applications of edge AI. In smart agriculture, the cameras mounted on tractors, robots, and drones maximize the efficiency of planting, fertilizing, and harvesting of the crops with the help of edge analytics. Similarly, the edge AI system analyzes multiple video data to improve farming operations and assess how farming practice impacts productivity.

Although these use cases fall into the multi-camera AI, these use cases are distinguished from the use cases in Section 3.3 in the sense that all the cameras are not deployed for the same purpose. For example, in a smart shopping cart, one camera is used for barcode reading and other cameras are for detection and localization, for which various AI and vision processes are running simultaneously. The AM69 processor is more effective in such use cases thanks tot he heterogeneous architecture.

4 Software Tools and Support

While being such a powerful processor, the AI application programing on AM69A is made simpler and faster with the Processor SDK Linux for AM6xA (PSDK Linux)[(4)], which provides the building blocks for customers to start developing smart-camera applications. The PSDK Linux leverages and enables an interplay of multiple open-source components such as GStreamer, OpenVX, OpenCV, and deep learning runtime such as TFLite, ONNX, and TVM on top of the foundational Linux® component and the firmware packages for remote cores and hardware accelerators. This makes edge AI application development on the AM69A as easy as programming in Python® or C++ while still taking advantage of hardware accelerators for vision processing and AI functions. Furthermore, since the PSDK Linux is designed to provide the unified software development framework for the AM6xA scalable processor family, the development experience on the AM62A or AM68A device makes the development on the AM69A much easier. The PSDK Linux documentation has more details.

The reference applications in PSDK Linux showcase perception-based examples such as image classification, object detection and semantic segmentation in both Python and C++ variants. TI also has converted and exported 100+ models from their original training frameworks in PyTorch, TensorFlow, and MXNet into a format friendly to the C7xMMA architecture and hosts them in the Edge AI Model Zoo[(3)]. In this process TI makes sure that these models provide the optimized inference performance on TI's embedded processors. The reference applications and the optimized models provide a good starting point for customers to explore high-performance deep learning on AM69A.

TI also provides Edge AI Studio that is a collection of tools to accelerate the development of edge AI applications on TI's embedded processors including AM69A. Edge AI Studio allows building, evaluation, and deployment of deep learning models. There are two tools as a part of Edge AI Studio: model analyzer and model composer. Model analyzer allows connection remotely to a real evaluation hardware to deploy and test AI model performance on TI's embedded processors. Model analyzer helps a lot during the model evaluation phase. Model composer is a fully-integrated solution for creating Edge AI applications, which helps collect, annotate data, train, optimize, and compile AI models for TI's embedded processors. Model composer enables the retraining of the models from Edge AI Model Zoo to fine tune the performance for the unique application requirements with custom data.

5 Conclusion

The AM69A processor is the best solution for the development of various Edge AI applications with as many as 12 cameras in terms of the performance and the ease of development. The heterogeneous architecture with multiple cores and hardware accelerators provides flexible ways to optimize the performance. Furthermore, the PSDK Linux with Edge AI Studio, and Model Zoo streamline the development process including training, optimizing, and deploying of the model to the application deployment with GStreamer, OpenVX, and OpenCV.

The AM69A is ready for the developer to start developing Edge AI applications with the AM69A starter kit EVM and the Processor SDK Linux for AM69A. Documentation, prebuilt images, and demonstrations are available from the AM69A product page. Find more information about Edge AI across TI's processors from TI’s Edge AI folder.

6 References

1. Texas Instruments, *Edge AI Smart Cameras Using Energy Efficient AM62A Processors Technical White Paper*
2. Texas Instruments, *Advanced AI Vision Processing Using AM68A for Industrial Smart Camera Applications Technical White Paper*
3. Texas Instruments, Edge AI Model Zoo
4. Texas Instruments, Processor SDK Linux for AM69A

IMPORTANT NOTICE AND DISCLAIMER

Mailing Address: Texas Instruments, Post Office Box 655303, Dallas, Texas 75265