INL/RPT-23-74920

Light Water Reactor Sustainability Program

Considerations regarding the Use of Computer Vision Machine Learning in Safety-Related or Risk-Significant Applications in Nuclear Power Plants

September 2021

U.S. Department of Energy

Office of Nuclear Energy

INL/RPT-23-74920

Considerations regarding the Use of Computer Vision Machine Learning in Safety-Related or Risk-Significant Applications in Nuclear Power Plants

Ahmad Al Rashdan[1], Roman Shaffer[2], Edward (Ted) L. Quinn[2], Michael Bailey[2], Ronald Jarrett[2], Kellen M. Giraud[1]

[1] Idaho National Laboratory
[2] Technology Resources

September 2021

Prepared for the
U.S. Department of Energy
Office of Nuclear Energy

ABSTRACT

With the advancements made to date in the field of artificial intelligence (AI), significant potential exists to utilize AI capabilities for nuclear power plant (NPP) applications. AI can replicate human decision making and it is usually faster and more accurate than humans. For implementations that impact critical NPP applications (e.g., safety-related or non-safety systems that potentially affect overall plant risk), a deeper safety analysis of the AI methods is necessary. AI applied to NPP operations could resemble the use of digital I&C (DI&C) because such applications involve digital computer hardware and custom-designed software that input plant data, execute complex software algorithms, and output the results to a system or licensed human operator to potentially provoke an action. For AI methods to be compliant with current safety requirements for DI&C, AI compatibility must be evaluated, and AI-related gaps may exist that prevent the prompt deployment of AI in NPPs. This effort aims to evaluate how example AI technologies align with the DI&C safety framework, and discusses how they could be analyzed, modeled, tested, and validated in a manner similar to typical DI&C technologies.

Because AI is a broad field that encompasses areas such as machine learning (ML), natural language processing, and computer vision, this research focused on a subset of methods categorized as the computer vision ML (CVML) methods. This report explores two CVML use cases, gauge reading and fire watch, considered relevant to the DI&C standards, as they could play a safety-critical role. For the gauge reading use case, a CVML-enabled technology that can read gauges at oblique angles is utilized. For the fire watch use case, a CVML-enabled technology is utilized that migrates fire watch from a manual (human) approach to automated fire detection. These use cases are mainly intended to give context to the CVML system discussion.

This effort assumes the worst-case scenario, with the CVML system being used to replace a safety-related or risk-significant system, thus requiring evaluation. Evaluating CVML against most of the relevant safety requirements for DI&C yielded several CVML-specific considerations due to the uniqueness of its characteristics in comparison with typical DI&C systems. For example, CVML models often employ commonly used (open-source) datasets, and it is not always possible to determine the level of overlap among open-source datasets. Therefore, the independence of the developed CVML models when demonstrating diversity is questionable, therefore creating vulnerability to common cause failure (CCF). The design verification process is also impacted since the data overlap could result in overestimation of the software validation and verification (V&V) performance results. Section 2 of this report evaluates a list of the identified CVML-specific characteristics and discusses the resulting considerations and potential solutions in the context of each referenced requirement. A summation is provided in Section 3.

This report is not to be used as a guideline. It was developed to identify and consider issues in the implementation of ML technologies used to augment activities that may have a bearing on plant operation. The report draws parallels to the use of DI&C technologies, for which many standards are available to guide their use in nuclear plant operation. It considers the technologies and some of the

potential implications of their use in safety-related applications but is not intended to address regulatory or licensing related issues.

Acknowledgements

The authors wish to thank the Light Water Reactor Sustainability (LWRS) program for funding this effort. They also wish to thank Curtis Smith, Roger Boza, Yugandhar Police, and Michael Griffel for their input, and John Shaver for his technical review and edits.

CONTENTS

FIGURES

TABLES

ACRONYMS

AGRS	Analog gauges reading system
AI	Artificial intelligence
ANN	Artificial neural network
ASME	American Society of Mechanical Engineers
BTP	Branch Technical Position
CCF	Common cause failure
CDF	Core damage frequency
CGD	Commercial-grade dedication
CGI	Commercial-grade item
CM	Configuration management
CNN	Convolutional neural network
CPU	Central processing unit
CV	Computer vision
CVML	Computer vision machine learning
DBE	Design-basis event
DCNN	Deep convolutional neural network
DI&C	Digital Instrumentation and control
D3	Diversity and defense-in-depth
EPRI	Electric Power Research Institute
ESPS	Engineered Safety Protection System
FMEA	Failure Modes and Effects Analysis
FPP	Fire protection program
FSAR	Final safety analysis report
GDC	General Design Criterion
GPU	Graphics processing units
HFE	Human factors engineering
HSI	Human-system interface
I&C	Instrumentation and controls
IEEE	Institute of Electrical and Electronics Engineers
LAR	License amendment request
LSTM	Long short-term memory
LWR	Light-water reactor
LWRS	Light Water Reactor Sustainability

ML	Machine learning
NEI	Nuclear Energy Institute
NFPA	National Fire Protection Association
NN	Neural network
NPP	Nuclear power plant
NRC	Nuclear Regulatory Commission
NUREG	nuclear regulatory
PRA	Probabilistic risk assessment
QA	Quality assurance
RG	Regulatory Guide
RNN	Recurrent neural network
SIL	Software integrity level
SLC	Standard license condition
SQAP	Software quality assurance plan
SRP	Standard Review Plan
TS	Technical specifications
V&V	Validation and verification

CONSIDERATIONS REGARDING THE USE OF COMPUTER VISION MACHINE LEARNING IN SAFETY-RELATED OR RISK-SIGNIFICANT APPLICATIONS IN NUCLEAR POWER PLANTS

1. INTRODUCTION

A key tenet of plant modernization and economic viability for the operating U.S. light-water reactor (LWR) fleet is that of reducing the high labor requirements for operating and maintaining nuclear power plants (NPPs), as driven by the current plant technology.

Since its inception, the U.S. Department of Energy Light Water Reactor Sustainability (LWRS) program has conducted R&D activities over a broad range of NPP functional areas, including addressing critical issues of technology obsolescence, plant reliability, plant worker efficiency, and operating and maintenance cost reduction [1]. This R&D is targeted to ensure that the NPPs are positioned for sustained operation in support of the national goals of energy security and environmental objectives [2]. Development and demonstration of these technologies and related methodologies have been conducted with collaborating partners (e.g., nuclear utilities, nuclear industry suppliers, and other research organizations) who share the same sustainability objectives. The result is a set of proven technologies that collectively address the requirements for the much-needed modernization of legacy plant systems—along with related operations and support processes—to ensure long-term sustainability and economic viability. This includes extending these technology types into areas that are currently challenging, as well as combining them with advanced analytics, artificial intelligence (AI) methods, and risk assessment tools—and in some cases, defining entirely new types of technologies to address plant operation and maintenance and to support requirements even more effectively.

With the advancements in AI to date, the potential exists to utilize AI capabilities in NPP applications [3]. While such applications have mostly been research-oriented, AI continues to evolve, having now perhaps reached a level of advancement that warrants broader consideration for certain applications in NPP environments. AI can replicate human decision making and is usually faster and more accurate than humans. For implementations that impact critical NPP applications, including safety-related or non-safety systems that potentially affect overall plant risk (i.e., core damage frequency [CDF] or large early release frequency), a deeper analysis of these AI methods is necessary to enable the responsible authority to make informed decisions that help ensure public and environmental health and safety. This is why a critical aspect of modernizing the U.S. operating LWR fleet is the regulatory oversight conducted by the U.S. Nuclear Regulatory Commission (NRC). As the need for modernizing the LWR fleet is recognized and actions are taken, there is a concomitant need to evaluate the technologies' ability to align with NRC's regulatory framework so that the advanced technologies being proposed for the modernization effort can be effectively regulated. This will ensure that the industry's high safety standards are maintained, and that they are synchronized with the program innovation in the pursuit of more efficient NPP operations.

AI used in operating NPPs could resemble the use of digital instrumentation and control (DI&C), because AI applications involve digital computer hardware and custom-designed software to input plant data (e.g., process or environmental data), execute complex software algorithms, and output the results to a system or licensed human operator to potentially provoke an action. Generally, DI&C systems fall into one of two categories: safety-related or non-safety-related. Safety-related systems are relied on to mitigate the consequences of accidents (i.e., design-basis events [DBEs]) by preventing core damage or offsite radiation release in excess of the limits given in 10 CFR Part 100 [4]. Safety-related systems can be protection systems (e.g., the reactor trip system) or an engineered safety feature system that ensures adequate core cooling during an event (e.g., valve actuation or emergency cooling pump start). Other

safety-related systems can support protection or engineered safety features, or entail a level of risk-significance that requires them to be designed, developed, and maintained to the relevant safety standards. Safety-related systems are required to be developed, design-verified, controlled, and maintained under an approved quality assurance program that meets the requirements of 10 CFR Part 50, Appendix B [5]. This includes both the hardware and software used in—or by—the safety-related system.

NPP systems that are not needed to prevent or mitigate DBEs are categorized as non-safety-related. These systems are not required to meet any specific regulatory requirements other than non-interference with the design functions of safety-related systems. However, some non-safety systems could be relied on to perform certain functions that carry a degree of risk-significance. In the context of NPP risk, if the plant's probabilistic risk assessment (PRA) determines that a non-safety system contributes to preventing or mitigating a DBE, mitigating the effects of beyond DBEs (e.g., anticipated transient without scram, station blackout), reducing the risk of a DBE (i.e., the cumulative risk effects), or significantly enhancing the diversity and defense-in-depth (D3) of the plant, the non-safety-related system could receive special treatment. The full regime of safety-related requirements may not be imposed on such systems, but an augmented quality approach may be taken to verify that the minimal acceptable quality standards are applied to ensure that the non-safety-related system can perform its design functions and meet the NPP's PRA goals.

1.1 Objective

Depending on the targeted applications and their safety classifications, AI methods must comply with current safety requirements for DI&C, and AI gaps might exist that prevent prompt deployment in NPPs. As such, acceptable safety-related AI systems for use in U.S. NPPs must be designed, fabricated, installed, and tested to quality standards commensurate with the importance of the safety functions to be performed. Safe-use issues would need to be addressed prior to actual implementation of the AI technology at an NPP.

This report evaluates a subset of AI technologies (i.e., computer vision machine learning [CVML]) in light of the safety framework, discussing how these technologies could be analyzed, modeled, tested, and validated in a manner similar to typical DI&C technologies. This effort is not intended to serve as a guideline for safety-use or regulatory compliance; rather, it is meant as an introduction to the safety bases for DI&C and associated considerations for deployment in a NPP as applied to a subset of AI technologies.

1.2 Scope

Because AI is a broad field that includes areas such as machine learning (ML), natural language processing, and computer vision (CV), the intent is to focus this research on a specific subset of methods categorized as CVML methods (discussed in Section 1.4.2). CV is, in general, a scientific discipline for extracting information from images (e.g., video sequences and photos). From an ML perspective, CVML is the process of isolating and extracting features from images, then making decisions based on those features. Common fields of application include object detection, image and object classification, and object localization or segmentation.

CVML utilizes complex models[a] coded into software algorithms. The software is loaded into a dedicated hardware and interfaces with various hardware sensors representing a DI&C system. As was discussed, for non-safety-related applications without risk-significance, only the relevant guidance need be applied to ensure that the minimum acceptable quality requirements are satisfied for implementation in an NPP. However, this report explores two use cases (i.e., gauge reading and fire watch) considered relevant to the safety framework, as they could play a safety-critical role and require a license amendment

[a] An ML model is a mathematical function that represents the relationship between the input and output.

request (LAR). Gauge reading and fire watch (discussed in the following sections) require many plant staff hours per day and are easily automatable from a technical perspective.

The use cases in this report mainly serve to provide context for the report's evaluation of CVML against safe-use criteria for DI&C systems, and to highlight important considerations and their potential solutions. Furthermore, the requirements discussed in this report do not provide a comprehensive or complete list of all requirements that must be considered within the DI&C safety framework. The report only applies the relevant safe-use criteria, selected in accordance with their potential degree of impact on the CVML development process, in alignment with the research objective of this report.[b]

1.2.1 Gauge Reading

As the LWR fleet ages, the NPPs have difficulty maintaining safe operations while using increasingly obsolete equipment. NPPs are outfitted with several thousand analog instruments. They are modernizing, and digitization is replacing certain analog I&C, but the slowness of this progress is expected to impact long-term plans for NPPs. In the short- to medium-term, plants will continue using analog instrumentation, including analog gauges. Many NPP safety-related systems still utilize analog gauges in the control room and throughout the machinery spaces. As part of normal operations, operator logs on these gauges must be maintained—itself a time-consuming process. Combining aging equipment with the potential for human error leads to the conclusion that the ability to automate the reading of analog gauges may generate significant cost savings and reduce human error. CVML technologies could assist in the reading and auditing of analog gauges, thus augmenting or replacing the manual gauge reading and logging process in NPPs.

CVML can analyze an analog gauge image to determine the measured parameter value, then input that measurement into a digital system that evaluates parameters against specific acceptance criteria. Implementation can take various forms (see Figure 1), including a system mounted in the control room to read indication gauges on the main control board. Another form of implementation is to place a system in remote plant locations to read analog gauges that plant operators must manually read during periodic operator rounds. The data are acquired using a set of video cameras positioned in strategic locations, arranged to allow for the images to primarily be straight on, thus minimizing their angle. However, because achieving a straight-on view of all the gauges in a control room panel is challenging, a method of reading gauges at oblique angles is needed. The LWRS program created a CVML-enabled technology that can read gauges at oblique angles using a method with a high success rate.

The gauge reading use case targeted by this effort assumes that the CVML system[c], called the analog gauges reading system (AGRS) in this use case, is being used to monitor safety-critical analog gauges (i.e., ones that are safety-related or risk-significant and must therefore be evaluated from a regulatory perspective). This use case assumes that AGRS is being used for monitoring and that it is not involved in actuating safety shutdown equipment. It also assumes that the envisioned use of the CVML system includes ensuring compliance with plant technical specifications (TS). For example, AGRS could be installed to monitor analog gauges on the Engineered Safety Protection System (ESPS). The plant operators are then required by the TS to periodically review the analog gauges to ensure that all channels of instrumentation fall within an acceptable indication range. Manual review of gauges can be replaced through installation of the AGRS. The AGRS cameras would be installed in the Main Control Room and would face the analog gauges used to satisfy the TS surveillance requirements for periodic ESPS channel checks. AGRS will need to satisfy the periodic channel checks in the ESPS TS and would require NRC approval before modifying the TS basis for compliance with the surveillance requirements. To obtain

[b] This report is not to be used as a licensing guideline and was developed solely for research purposes. It entails the usual subjectivity and potential shortfalls.

[c] The term *CVML system* is used in this report to describe the software (e.g., the CVML model and algorithm) and the hardware that implements the software and connects to other related hardware to achieve the function being automated.

NRC approval, an LAR must be submitted[d]. Because AGRS is based on digital technology/software, the LAR must meet DI&C-ISG-06 [6] guidance for licensing DI&C systems. Section 2 discusses example requirements that would have to be addressed for an NPP-specific AGRS LAR submission.

Figure 1. Automated gauge reading enables any camera to be used to log several gauges autonomously.

1.2.2 Fire Watch

Fire protection programs (FPPs) at U.S. NPPs aim to minimize both the probability of occurrence and the consequences of fire. To meet these objectives, FPPs for operating NPPs are designed to provide reasonable assurance—in part, through defense-in-depth measures—that a fire will not prevent the necessary safe shutdown functions, and that radioactive release to the environment in the event of a fire will be minimized. These defense-in-depth principles are aimed at achieving the following objectives:

- Prevent fires from starting
- Rapidly detect, control, and promptly extinguish any fires that do occur
- Provide protection for structures, systems, and components important to safety, so that any fire not promptly extinguished by the fire suppression activities will not prevent safe shutdown of the plant.

Fire watch, a visually demanding fire protection function, may need to be performed frequently to meet the requirements of the NPP's FPP. Fire watches typically involve periodic inspection of an area to ensure that no fire is starting or in progress. It is needed to compensate for temporary fire protection/suppression system outages (e.g., for maintenance or component repair), or when performing fire-hazardous activities. The length of time between each inspection is based on the fire risk and the

[d] One option for fleetwide deployment of the CVML system is to submit to NRC a topical report for AGRS. The approved topical report will provide a justification for major aspects of the DI&C licensing requirements addressed by Licensing Process Interim Staff Guidance (DI&C-ISG-06) Revision 2. This would simplify the LAR process for NPPs implementing the technology, as the LARs would only need to address the plant-specific action items identified in the topical report to credit the AGRS system as the basis for satisfying manual surveillance checks as per the TS.

equipment in the area. A continuous fire watch is the most strenuous in terms of manual actions. However, it may not require that a person always remain in the area, but rather that the area is inspected every 15 minutes. An hourly fire watch requires a person to inspect the area every hour. For some fire watches, inspections are to occur every six hours or once per shift (12 hours).

The LWRS program created a CVML technology that can migrate fire watch from a manual (human) model to an automated one. It involves replacing human-based fire watch with a video camera and CVML-based fire/smoke detection model and algorithm (Figure 2). Initial implementation of the CVML system for fire watch could take the form of a mobile system (e.g., a crash cart or suitcase device) that can be relatively easily deployed in the NPP environment to take the place of fire watches conducted by human operators.

Figure 2. Automating fire detection via a video stream enables a camera and CVML system to perform fire watch.

Regulatory considerations for the FPP differ from those pertaining to the previous gauge reading use case. NPPs entail a standard license condition (SLC) under which the FPP is controlled. In accordance with 10 CFR Part 50.48(a) [7], each operating NPP must have an FPP that satisfies General Design Criterion (GDC) 3 ("Fire Protection") of Appendix A [8], "General Design Criteria for Nuclear Power Plants," to 10 CFR Part 50, "Domestic Licensing of Production and Utilization Facilities" [9]. In addition, plants that were licensed to operate before January 1, 1979, must meet the requirements of 10 CFR Part 50, Appendix R, "Fire Protection Program for Nuclear Power Facilities Operating Prior to January 1, 1979," [10] except to the extent provided for in 10 CFR Part 50.48(b). Plants licensed to operate after January 1, 1979, must comply with 10 CFR Part 50.48(a), as well as any plant-specific fire protection license condition and TS.

There are two approaches to fire protection in NPPs. The deterministic FPP approach means that the NPP is subject to 10 CFR Part 50.48(b) and Appendix R. This is the most conservative approach. NPPs that are subject to Appendix R and entail an SLC are required to perform safety evaluations of any changes to the FPP. The implementation of this approach would have to ensure that no impact to the FPP results in a deviation from the commitments made under the SLC. For cases in which these requirements are impractical, or when there is a more favorable approach for achieving an equivalent level of safety, NRC regulations allow for exemptions from specific requirements of the regulations when certain enumerated special circumstances exist (10 CFR Part 50.12 [11]). NRC will grant such exemptions only if they do not present an undue risk to public health and safety, as well as meet other requirements.

Each FPP change would be evaluated and, depending on the results, the NPP could implement the change without prior NRC approval (e.g., as an update to the Technical Requirements Manual—equivalent to an update under 10 CFR Part 50.59 [12]). Implementing the changes without prior NRC approval is only allowed if those changes do not adversely affect the ability to achieve and maintain safe shutdown in the event of a fire, as documented in a safety evaluation. The licensee should maintain, in auditable form, a current record of all such changes, including an analysis of the effects of the change on the FPP, and should make such records available to NRC inspectors upon request. If the licensee incorporated the FPP in the final safety analysis report (FSAR), all changes to the approved program should be reported, along with the FSAR revisions required by 10 CFR Part 50.71(e) [13]. Additional information can be found in Regulatory Guide (RG) 1.189 [14], "Fire Protection for Operating Nuclear Power Plants."

The risk-informed approach uses 10 CFR Part 50.48(c), adopted by NRC in 2004. It incorporates National Fire Protection Association (NFPA) 805 [15] by reference (with certain exceptions) and allows licensees to voluntarily adopt and maintain FPPs that meet the requirements of NFPA 805, instead of meeting the requirements of 10 CFR Part 50.48(b) or the plant-specific fire protection license conditions. In September 2005, the Nuclear Energy Institute (NEI) published industry guidance in NEI 04-02, Revision 1 [16]. RG 1.205 [17], "Risk-Informed, Performance-Based Fire Protection for Existing Light-Water Nuclear Power Plants," endorsed by NEI 04-02, offered additional information and guidance to supplement the NEI document and assist licensees in meeting NRC's requirements regarding fire protection at NPPs.

The risk-informed approach considers risk insights (as well as other factors) down to the individual component level to better focus attention and resources on design and operational issues, per their risk importance. This approach relies on a required outcome rather than requiring a specific process or technique to achieve that outcome. NFPA 805 enables leveraging the state of the art in fire protection evaluation techniques to maintain and enhance safety. The fundamental FPP uses the fire protection elements of prevention, fire detection and suppression, and safe shutdown. It allows nuclear safety performance criteria to be satisfied by using both fire modeling and quantitative fire risk evaluation to supplement the approach of deterministically failing all equipment within the fire area in order to evaluate the capability for safe shutdown. By using the risk-informed approach, resources can be focused on higher risk areas.

The use cases targeted by this report assume an NPP subject to Appendix R. The NFPA 805-based approach is not evaluated herein because it relies on fire modeling and quantitative fire risk evaluation and could ultimately reduce the fire watch requirement to the point where automation is not essential. For single deployments of a CVML system (or less than a nominal number) at an NPP subject to Appendix R (e.g., a crash cart or suitcase device), the safety evaluation may be relatively straightforward and would likely not require prior NRC approval (i.e., it would not fall under 10 CFR Part 50, Appendix B unless the NPP made such commitments in its licensing basis or SLC). A safety evaluation must still remain on file for NRC auditing, as well as to address the typical requirements (including the ones discussed in Section 2) via an appropriate level of rigor. This level is commensurate with the impact on the safety-related protection systems, thus ensuring that NRC expectations for changes are met, without prior NRC approval under the SLC. Deploying multiple CVML systems may impact the safety-related protection systems' ability to maintain safe shutdown of the NPP, in which case the safety evaluation would be more rigorous and likely require prior NRC approval. This is also the case when a proposed change involves altering a license condition or TS used to satisfy NRC requirements. In cases where the CVML technology does make an impact, the safety evaluation should address all applicable regulatory requirements. An LAR will be submitted for NRC review (unless the NPP is granted an exemption). Such evaluations might not be conducted via the same level of rigor as for the safety-related DI&C systems, but must nonetheless ensure adequate justification that the CVML system does not impair the safety-related protection systems.

In this effort, the fire watch use case assumes that the criteria used are the same as for safety-related systems (using Section 2). This assumption is used to highlight the CVML system's applicability and gaps, as well as potential solutions associated with CVML.

1.3 Safety Framework for Digital Instrumentation and Control

A background in regulations, guidelines, and standards for DI&C systems within a NPP is required to understand safety considerations for CVML systems. Thus, this section summarizes the safety framework for DI&C.

1.3.1 Requirements

The governing framework for safety-related DI&C in NPPs is 10 CFR, Part 50, "Domestic Licensing of Production and Utilization Facilities," specifically:

- 10 CFR Part 50.55a(h), "Protection and safety systems" [18]
- GDC 21, "Protection System Reliability and Testability," of Appendix A, "General Design Criteria for Nuclear Power Plants" to 10 CFR Part 50
- Criterion III ("Design Control") of Appendix B, "Quality Assurance Criteria for Nuclear Power Plants and Fuel Reprocessing Plants" to 10 CFR Part 50.

The regulations in 10 CFR Part 50.55a(h) require that NPP protection systems adhere to Institute of Electrical and Electronics Engineers (IEEE) Std. 603-1991, "Standard Criteria for Safety Systems for Nuclear Power Generating Stations" [19], along with the correction sheet dated January 30, 1995. With respect to the use of computers in safety systems, IEEE Std. 7-4.3.2 2003 [20], "Standard Criteria for Digital Computers in Safety Systems of Nuclear Power Generating Stations," provides computer-specific requirements to supplement the criteria and requirements of IEEE Std. 603-1998 [21], "Standard Criteria for Safety Systems for Nuclear Power Generating Stations."

The design requirements of 10 CFR Part 50, including the need for redundancy, diversity, and defense in depth, are based on the need to ensure reliable system functionality in the face of a wide range of failure modes—up to and including the design-basis accidents described in each NPP's updated FSAR, and in the combined operating license and design certification applicants' FSARs.

One requirement in GDC 21 is that protection systems (or safety systems) be designed for high functional reliability commensurate with the safety functions to be performed. In part, Criterion III requires that licensees specify quality standards and provide design control measures for verifying or checking the adequacy of safety system designs.

1.3.2 Graded Approach

As mentioned earlier, a non-safety-related system is given special treatment if it performs or supports the performance of a risk-significant function, such as:

1. Responding to a beyond DBE (e.g., anticipated transient without scram, station blackout)
2. Ensuring long-term NPP safety after an event
3. Minimizing CDF and large early release frequency.

CVML systems may be used by NPP operators to perform risk-significant actions related to the above functions. In such cases, a graded approach could be appropriate. Guidance related to a graded approach for non-safety systems is found in IEEE Standard 1012-2004, "IEEE Standard for Software Verification and Validation," [22] endorsed in RG 1.168, "Verification, Validation, Reviews, and Audits for Digital Computer Software Used in Safety Systems of Nuclear Power Plants" [23].

The IEEE standard identifies four software integrity levels (SILs) (i.e., 1–4, with SIL4 being equivalent to a nuclear safety-related system under the full purview of NRC). Using IEEE 1012, the appropriate SIL rating would be assigned to the CVML system, based on the likelihood of failure and the consequence to the plant and personnel (and environment). The SIL rating would dictate the quality requirements imposed on the system at a level of rigor commensurate with the SIL rating, based on the criteria in the IEEE standard. A justification for using the graded approach should be provided, along with the basis for the assigned SIL rating and a description of the graded approach used in developing the CVML system. The applicability of the review criteria in Section 2 would be determined on a case-by-case basis. This includes the appropriate level of detail per the SIL rating.

1.3.3 License Amendment

After the initial operating license is granted to the licensee, the license may be amended, renewed, transferred, or otherwise modified depending on activities that affect the reactor throughout its operating life. Review of LAR applications is a primary mechanism for regulating changes in NPP operations. The proposed change to licensed NPPs could be subject to an LAR review by NRC staff prior to implementation. Approval or denial of LAR applications is part of the continuous process of managing issues related to NPPs.

1.3.3.1 License Amendment Request Content

The docketed LAR should contain sufficient necessary information to demonstrate regulatory compliance. Such information can be derived from a variety of concept, system requirement, hardware or software requirement, or design documents (including the ones discussed in Section 2). Along with the system design, the development process is also included in the LAR to support NRC's determination that the design meets regulatory requirements and that, in NPP safety-related applications, the process is of sufficiently high quality to produce suitable systems and software.

The level of information and detail included in the LAR to demonstrate compliance with regulatory requirements may vary in accordance with design, configuration, and operational features unique to the proposed digital modification. As part of the NRC oversight process, NRC staff may perform one or more inspections to evaluate compliance with the license conditions related to the license amendment.

1.3.3.2 Review Guidance

Review guidance is provided by NRC staff to clarify the regulatory requirements of 10 CFR. Guidance comes in various forms, such as RGs and nuclear regulatory documents (NUREGs), and can be used by NPP personnel to meet NRC requirements for DI&C. Generally, RGs describe methods that NRC staff consider acceptable for meeting the requirements contained in 10 CFR Part 50. On its own, an RG is not considered a requirement unless the NPP commits to it during regulatory review. RGs to which the NPP commits during regulatory review become part of the NPP's licensing basis.

Reviews of DI&C equipment modifications focus on the system architecture design, human-system interfaces (HSIs), and hardware/software architectures in order to determine whether the four fundamental design principles—redundancy, independence, deterministic behavior (i.e., predictability and repeatability), and D3 —have been met to ensure that the design affords adequate safety. These principles form the basis for many of the regulatory requirements and associated regulatory guidance that stem from NRC and industry consensus standards (e.g., IEEE) and are referenced throughout this report.

IEEE Standard (Std) 603-1991, "IEEE Standard Criteria for Safety Systems for Nuclear Power Generating Stations," and IEEE Std 7-4.3.2-2003, "IEEE Standard Criteria for Programmable Digital Devices in Safety Systems of Nuclear Power Generating Stations," as endorsed by RG 1.152 [24], "Criteria for Use of Computers in Safety Systems of Nuclear Power Plants," are used to review the proposed DI&C equipment modification, in accordance with applicable regulatory requirements and the plant's licensing basis.

Detailed I&C guidance is contained in NUREG-0800[e] [25], "Standard Review Plan for the Review of Safety Analysis Reports for Nuclear Power Plants: LWR Edition," Chapter 7, "Instrumentation and Controls." When a license amendment is required, the LAR should describe the functions of I&C equipment identified in the FSAR, as updated, along with the equipment to implement those functions. Additionally, the LAR should identify the parts of the licensing basis that are being updated as a result of the proposed change. The LAR should ensure that the provided documentation demonstrates regulatory compliance of the safety-related system. Chapter 7 of Appendix 7.0-A [26], "Review Process for Digital Instrumentation and Control Systems," and Branch Technical Position (BTP) 7-14, "Guidance on Software Reviews for Digital Computer-Based Instrumentation and Control Systems," [27] provide insight into the review of digital systems in support of safety evaluations.

Before discussing the research targeted by this scope, it is necessary to introduce some ML and CVML basics that would be of benefit to that discussion. This introduction is for readers with no background in ML, and can be skipped for readers knowledgeable on this topic.

1.4 Machine Learning Basics

ML is an artificial construct that attempts to mimic the human learning process; hence, it is considered a subset of the field of AI. In ML, algorithms learn from and make predictions based on mathematical models built on input data, and through these algorithms, the computer learns how to complete a given task with increasing efficiency and accuracy [28,29]. In ML, the selected algorithm (e.g., an artificial neural network [ANN]) must be trained to perform its intended task (e.g., classification to perform object identification). During the training process, the algorithm continuously adjusts the weights of the mathematical model to produce accepted or accurate predictions at the output(s). When the correct behavior reaches the threshold of acceptance for a specified measure, the model is ready for real-world application. Training involves multiple steps and therefore requires multiple datasets for training and validation and a testing (hold-out) dataset for formal testing to confirm acceptable operation.

As the algorithm is trained, its prediction accuracy improves with each pass[f] of the data through the algorithm (in this case, an ANN). The validation data can be used to determine an appropriate stopping point[g] for the training process, such as an error function or accuracy measure. The test dataset is comprised of structurally similar data that have not been used for training or validation. The algorithm's performance is compared to the test dataset and evaluated against the acceptance criteria. If the performance is satisfactory, the algorithm is deployed/commissioned for the target environment. The test data usually account for 10–20% of the total available dataset. The common approach is to allocate a larger portion of the remaining 80–90% of the dataset to training rather than validation (e.g., a 70/30 split), but other approaches use a larger portion for validation (e.g., a 20/80 split).

Figure 3 from [30] shows that ML can be performed through supervised, unsupervised, and reinforcement learning, and has various applications (e.g., classification, clustering, and regression) to general use cases (e.g., targeted marketing, game AI, weather forecasting, diagnostics). The definition of ML types can be described as:

Unsupervised Learning. This is, in a sense, an open-loop process whereby the data are input to the model and separated into similar groups (i.e., clustering), or the non-essential data are removed (i.e.,

[e] 10 CFR Part 50 references IEEE Std 603, which is for protection systems but was written before the widespread use of digital technology. To supplement the guidance, IEEE Std 7-4.3.2 was written to address digital technology and software. Both *603 and *7-4.3.2 were endorsed by NRC staff through RGs, which the staff use to explain in detail how to apply a particular document to meet NRC regulatory requirements. NUREG-0800 is specifically aimed at NRC staff, and explains how the staff is to perform regulatory review of a license application. This ensures consistency in NRC's work and provides the industry with clarity on what to expect during such detailed staff reviews.

[f] The training dataset can be separated into batches. An "epoch" results when all the batches have passed through the algorithm.

[g] When the prespecified number of epochs is reached, a loss function (basically, an error function between the expected output and predicted output) is calculated, and that loss function is used to determine whether testing can stop.

dimensionality reduction). This process requires neither a human to participate in the training process nor an example of correct output to perform the activity—other measures are used to determine success. Among other things, the model learns on its own how to categorize the data, based on hidden structures.

Supervised Learning. This requires an example of correct (or expected/desired) output (the feedback), along with the input data to properly perform the activity. Examples of correct input are called labeled data, and a human is generally required to perform the labeling. Extensive effort is usually invested in this area of ML due to the massive amounts of data needed to perform this activity successfully (i.e., accuracy measured against some threshold of acceptance). The most common application is image classification.

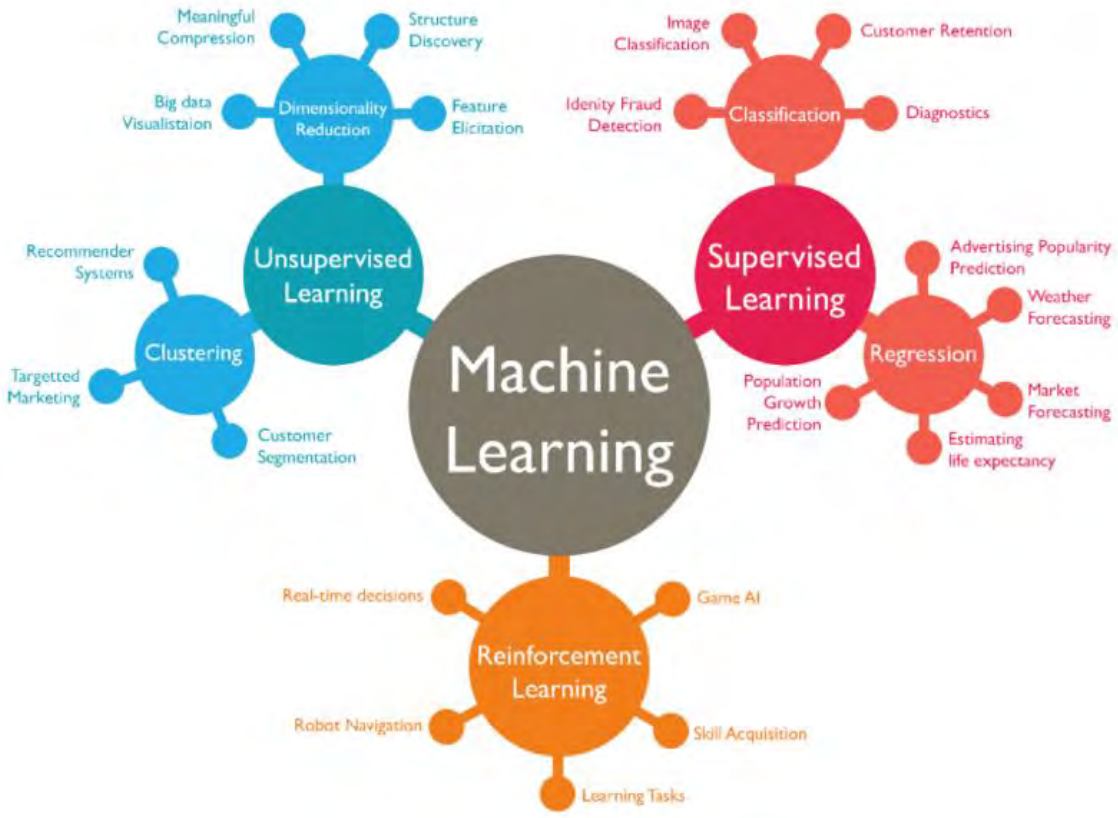

Figure 3. Machine learning taxonomy applications [30].

Semi-Supervised Learning. This is a combination of the previous two processes in that a small portion of the input data is labeled. For example, in using the labeled data as an input, the model learns the structure of the data, then predicts how to label (categorize) the unlabeled portion of the input data.

Reinforcement Learning. In this process, the algorithm learns the correct output on its own by receiving a "reward" (the feedback) that represents success in achieving the objective. Performing the task numerous times (e.g., thousands) allows the algorithm to determine the best method of achieving the objective.

One set of methods very widely used in AI/ML are ANNs, as they can be used in all applications of the above-listed learning paradigms. Given their importance to AI/ML, the remainder of this report describes them in further detail.

1.4.1 Artificial Neural Networks

A biological neural network (NN) is composed of neurons connected by synapses (from axons to dendrites). A neuron can be connected to many other neurons (i.e., one-to-many) in an extensive NN. An ANN emulates the biological NN by containing nodes that each process data and pass them to other nodes in a one-to-one or one-to-many network, depending on the network structure (which itself depends on the application). The powerful capability of ANNs to solve complex problems involving large amounts of data is the reason for the significant rise in their modern-day usage. An ANN is a generic term for various NN architectures across different applications. However, certain NN structures are better suited to solving certain problems:

Feedforward Neural Network. In a feedforward NN, the data pass through the different input nodes until reaching the output node. In other words, the data move in only one direction: from the input node(s) to the output node(s). Unlike in more complex NN types, there is no backpropagation. A feedforward NN may contain a single layer or have hidden layers (i.e., the layers between the inputs and outputs).

Deep Learning Neural Network. An NN architecture with multiple hidden layers is called a deep learning NN. In general, deep learning NNs perform better than other NNs but are harder to train and are computationally expensive to run. However, newer techniques and improved algorithms have led to widespread usage of deep learning NN techniques in ML implementations.

Convolutional Neural Network. A convolutional NN (CNN) uses a variation of the multilayer perceptrons. The CNN is a feedforward NN featuring a dedicated layer for performing the convolution function on the input data to extract features using a filter (weighted inputs). Different filters (with different weights) can extract different features from the input data. Following the convolution, a pooling function is used to eliminate unnecessary information but retain the essential feature information. Pooling is followed by an activation function that can extract non-linear features from the data. When a CNN contains multiple convolutional layers, it is referred to as a deep learning NN (or simply a deep CNN-DCNN). Fully connected (hidden) layers receive the output of the last convolutional layer, and these fully connected hidden layers interpret the feature representations and perform the high-level reasoning [31]. CNNs show very effective results in image and video processing, natural language processing, and recommender systems. They are also applied in signal processing and image classification, and for image analysis and recognition.

Recurrent Neural Network. A recurrent NN (RNN) is a type of ANN in which the output of a particular layer is saved and fed back to the input. This helps predict the outcome of the layer. The first layer is formed in the same manner as in the feedforward NN, namely, via the weighted sum of inputs. However, in subsequent layers, the feedback acts as a sort of memory (recurrent) such that, from one time step to the next, each node remembers some of the information it possessed in the previous time step. These nodes act as memory cells while computing and carrying out operations. The feedback values are also weighted and included in the sum of inputs that determines when the node fires. The RNN is effective where sequential information is important, which is why this type of NN is very effective in text-to-speech conversion technology.

Long Short-Term Memory Neural Network. The long short-term memory (LSTM) NN is a type of RNN that can carry forward temporal information for learning over time [32]. The LSTM NN controls the flow of information during learning, such that it can handle relatively long time lags between important features in time-series data, making it ideal for use with time-dependent data (e.g., video). This is achieved by using a memory cell that retains information during the time lag. But as the LSTM NN learns, a forget-gate allows for eliminating unnecessary information. LSTM NNs can solve complex problems that other RNNs cannot [33].

A review of the literature [3, 31, 34, 35, 36] reveals a wide variation in NN architectures—not only between the model categories (RNN versus CNN versus LSTM NN), but also within a single category

(e.g., CNN). Whereas the internal structures can be described and their functions explained in terms of how a node functions (e.g., the sum of weighted inputs reaches a specified threshold, causing the output to fire), as the complexity of the architecture grows, the ability to explain the internal workings of the architecture quickly diminishes. This is evident in [31, 35, and 36], among others, with multiple architectures being compared and contrasted in terms of performance measures imposed on the ANN outputs, as well as in terms of computational cost (e.g., hardware cost and time of execution) and training time. Furthermore, newer and better implementations are determined via iteration, which involves adding/removing layers, activation functions, etc., to determine which combinations improve on previous architectures regarded as benchmarks.

1.4.2 Computer Vision Machine Learning

Prior to advancements in ML, computer vision used manually extracted visual features (e.g., flames) and customized detection techniques. For example, color (e.g., RGB color content per pixel), shape, and motion properties were used to create a high-dimensional feature vector. This method required significant amounts of data and human-based engineering, making it difficult to train a decision-making process. With the advent of CNNs, deep learning, and LSTM NNs, CVML gained momentum as image analyzers, improving dramatically in terms of performance. A primary benefit of CVML is that the features are not manually engineered. Instead, the model is structured in a manner that allows it to extract what it determines to be the best features for a given application. CNNs—particularly DCNNs—are deemed the most efficient NN type for image analysis, as they introduce layers of kernels (i.e., filters) optimized to extract various sizes and forms of image features.

2. SAFETY EVALUATION OF CVML

This section introduces some key safety framework aspects for DI&C, then evaluates CVML's applicability to a specific set of safety requirements (extracted from standards and regulatory framework). The use cases discussed in Section 1.2 are used as context for demonstrating how the requirements apply to CVML. Within the applicability evaluation of CVML, issues or gaps are identified and potential solutions discussed (when available).

2.1 Design Control

The manufacturer (or engineer of record) of safety-related systems, components, and software must establish and implement a process to control the design of the safety-related product, including all design changes. The design process should address control over design inputs, outputs, changes, interfaces, records, and organizational interfaces pertaining to the organization and its suppliers. This ensures that design inputs are correctly translated to design outputs in sufficient detail to permit verification, which is to be performed by personnel not involved in the system design. 10 CFR Part 50, Appendix B, Criterion III requires, in part, that quality standards and design control measures for verifying or checking the adequacy of safety-related system designs should be used in the manufacturing of the safety-related system, components, and software.

From a design control perspective, it is important to control the process of design modification. The design and design change processes should be developed, documented, and revision controlled, including the roles and responsibilities of the personnel implementing the program. The modifications should be performed (controlled) in a manner similar to the original design, and be approved by the same design authority. One strength associated with CVML is its ability to constantly learn and improve model performance by continuously updating internal parameters. While periodic training of the CVML models may be no different than training for typical DI&C systems utilizing adjustable parameters during runtime (e.g., for system tuning), the CVML model could periodically adjust its internal parameters to improve output (e.g., data classification to determine "fire" or "no fire" for the fire watch use case), and the tuning is substantial enough to significantly alter the performance of the ML model. This presents a design change that must be controlled. The design change process for the CVML system should be detailed, including the change control process. The design change inputs (i.e., training dataset) should be carefully examined and filtered via the same metrics used in the initial design to ensure it matches the evaluation process and quality of the initial CVML design.

From a design input control perspective, many CVML models are derived from publicly available models. Open-source datasets have been historically important for advancing certain classes of CVML, given that massive amounts of data are needed for training and validation. While this has resulted in dramatic advancement of ML, it could present a challenge to satisfying design input control requirements. The original images/videos (used for training, validation, and testing) and models (used for feature extraction) should be placed under some form of configuration control (e.g., create a controlled local copy) to satisfy relevant safety requirements pertinent to configuration management (CM). This could include providing an approach to describe how the externally sourced data used to train, validate, and test the CVML system includes a means of control (i.e., ensuring the automated methods extract the same data every time). Additionally, tools must be developed to identify added images/videos and ensure that any revision of the data files is recognized before being used for model tuning.

From a design output control perceptive, modifying CVML models introduces another concern, since they often use probabilistic thresholding functions that could cause them to be sensitive to minor design decisions/changes and small variations in the operational image input, and thus potentially generate different output. This is especially apparent when overfitting occurs as a result of the ML model memorizing the training dataset too tightly (i.e., learning the structure of the training data, not the modeled process). This different output can also occur if the training data are too specific and

inadequately represent the actual data/environment. This could cause the CVML model output to become sometimes unrepeatable (i.e., unreproducible and unpredictable) due to even a minor design change. A sensitivity analysis should be performed to demonstrate the CVML's ability to generate consistent results after retraining with an identical set of images/videos, or following a design change. The means to achieve this remains a topic of ongoing research.

A QA program also requires identification and incorporation of appropriate design standards into the program, and that any deviations are adequately controlled. The procedures and processes should address the disposition of nonconforming items. In CVML, the quality of the training data is critical for enabling the model to perform the task for which it was intended. In this context, quality includes not only appropriateness for the intended application, but also trustworthiness of the source. For example, a well-known, commonly used data source in CV is YouTube-8M [34], which contains millions of videos, including thousands of annotations (labels). Other examples of CV open-source datasets are YFCC and FiSmo [32]. This large set of input data used in CVML is often extracted using automated image/video acquisition tools, without carefully examining the used input. The input data could include falsely annotated images that ultimately impact model performance. Therefore, the input data should be thoroughly vetted and documented for appropriateness to the application. The data must be examined (via metrics for measuring image quality and feature visibility) to ensure that it is of satisfactory quality. Validation by means of an independent testing dataset is another tool used to ensure that the performance is not impacted by the input data quality. However, validation is an indicator of the quality of the combined performance of the data and model, and ML models can sometimes make correct decisions for the wrong reasons. This is usually handled by explainability methods that reveal the CVML model's focus in decision making.

2.2 Design Verification

Design verification would ensure that changes are also subject to a design verification equivalent to—and conducted in the same manner as—the original design. This includes both design inputs and outputs. The extent of the design verification is a function of the item's safety significance, design complexity, and similarity to previously proven designs. A compliant QA program would include design verification processes to ensure that items, computer programs, and activities are suitable for their intended application, and are of a quality commensurate with their impact on safety.

Design verification procedures should be established and implemented to ensure that an appropriate verification method is used, the acceptance criteria are identified, and verification is satisfactorily accomplished and documented. Verification methods include design reviews, alternate calculations, and qualification testing. Testing should verify the acceptability of a specific design feature by demonstrating satisfactory performance under conditions that simulate the worst-case operational conditions expected in the system's intended installation location (e.g., adverse environmental conditions such as temperature, humidity, and radiation, as well as the operational load, including tasks related to network communications).

The design verification process should be developed, including how acceptance criteria are determined and measured. The selected performance measures for the CVML system should be appropriate for the intended application, and should accurately gauge the system's suitability for its intended purpose. It may be necessary to use several performance measures to determine the acceptability of the CVML. There are multiple such measures in the AI community, and the most frequently used ones are commonly benchmarked against each other [35, 37, 38]. The applied measure(s) are important for properly determining that the model not only reliably performs its intended task in the target environment, but also that it is appropriate for that task. The problem may be further compounded by skewed datasets wherein the desired class may be of rare frequency/distribution within the dataset [39]. The common measures are *accuracy*, *false positive*, *false negative*, and, for benchmarking between algorithms, *mean average precision*. Other measures include confusion tables, *precision*, *recall*, and *F-measure* [40].

The reasoning behind the major CVML design decisions should be discussed, such as the model architecture used, the size of the DCNN, the type of functions used within the layer, and all related design decisions, with evidence to support each decision made. Design characteristics (e.g., computational performance in terms of floating point operations per second [38], frames per second [38, 40], and the memory size of the algorithm itself [for applications targeting field programmable gate arrays and mobile devices]) [40] can be used to justify one design over another. Whatever the measures are, they should be documented along with the acceptance criteria to enable comparison of multiple designs (if applicable), as well as allow an objective review and determination of the adequacy of the overall design.

The CVML system use cases discussed in this report are based on DCNNs, which are particularly efficient as image classifiers. However, the black box approach means that improvements are often determined via experimentation [31, 34, 35, 36, 38]. LSTM NNs [33] were shown fairly recently to be promising image classifiers [32], but they feature the same black box functionality issue as those architectures with hidden layers to perform the high-level reasoning. Verification through design review is not always feasible since DCNNs often consist of a mathematical model that is incomprehensible to reviewers. The challenge of the black box testing is that it does not reveal the internal bias of the CVML models (for example, a CVML model trained with fire images that often show a fireman would start to flag the fireman as fire), and does not account for all case scenarios (e.g., gauge types with unique features, or fires of an unusual color), not to mention the worst-case scenarios (i.e., features that are very hard to detect, or degradation of the camera image quality as a result of temperature, humidity, or radiation). Therefore, in addition to black box testing, the inability to review the DCNN design should be compensated for by using explainable, transparent models to identify the "what" and "how" (not the "why"). In this sense, the "what" would be relatively simple decision making (yes/no or fire/no-fire), and the "how" would be measured by objective performance criteria for the success rate of predictions. The "why" would be to get inside the black box. Some implementations extract data from intermediate layers to achieve those objectives. For example, [41] described an architecture in which data were extracted from intermediate convolutional layers and further processed to isolate the feature(s) of interest *after* the high-level reasoning of the fully connected layers identified the target object(s) in the image.

Design verification should be performed by competent individuals or groups different from those who performed the original design. Ideally, the design verifiers should not be from the same organization (i.e., within the company). If this is infeasible, the design verifier(s) should not have been involved in selecting the design inputs, considerations, or approach, and should use a separate reporting line to the organization's quality organization in order to that ensure safety issues can be raised without fear of reprisal. Three types of independence should be developed: technical, managerial, and financial. At a minimum, in the case of safety-related applications of CVML technology, it should demonstrate that the quality organization was sufficiently independent from the design organization. Of particular importance are the qualifications and experience of the design verification staff (especially for software). The staff performing these functions should have the necessary knowledge and experience to thoroughly evaluate the CVML system design. This means that the person checking the design was not also involved in creating it. This might be feasible on the CVML application level, but the training of CVML models often utilized common training and validation data or feature extraction models (e.g., You Only Look Once). While this has resulted in dramatic advancements in all fields of AI, it could present a challenge to satisfying safety requirements. The diversity of the data and models (e.g., using different sets) should be demonstrated when testing models against defined acceptance criteria is sufficient to achieve acceptable performance when installed in the NPP target environment. Once used, the original source should be placed under configuration control to satisfy relevant safety requirements (e.g., design verification, CM, and defect reporting).

In the case of using commonly used feature extraction models, commercial-grade dedication (CGD), the process by which a commercial item is dedicated as a basic component for use in an NPP must be performed for those models (Section 2.6). This may pose a concern given the broader business model for

those models that likely did not include nuclear applications. Another consideration with verification is the acknowledged lack of repeatable results between developers due to the stochastic nature of the learning process. Even when the same person uses the same machine, two different training runs generate deviating results. This area of ongoing research is further discussed in Section 2.3.1.4.

2.3 Design Attributes

2.3.1 Architecture/Complexity

The new system's architecture must be developed and illustrated along with what changes, if any, must be made to the existing architecture to integrate the new system (including the reason for the changes). The documentation to describe the physical and functional architecture of the new system should be provided via both text and diagrams (e.g., functional/architecture block diagrams and functional logic diagrams). Sufficient information should be provided to understand what changes to the facility, if any, are necessary for installing the new digital equipment. The information should reflect the new system's architecture and design. The information should also define how the changes affect the existing plant (e.g., architectural restrictions on size, location, cooling, or power supplies).

The CVML system should be described along with its purpose, any necessary changes needing to be made to the facility in order to install the system, and interfaces with other systems and expected operator actions based on the CVML system output. The CVML system description should detail the software and related hardware architectures and how they interface with each other. To foster understanding, block diagrams should be used that show the software and hardware components, along with the interfaces with external systems.

The hardware architecture should describe each major subcomponent of the system:

- Input sensors (e.g., video cameras) and signal types or data format (e.g., frames per second, 1024 × 1024 pixels, visible light or infrared)
- Processing units and architecture (e.g., type and number of central processing units [CPUs]); for example, [31] describes a parallel computing architecture
- Resource requirements such as power/voltage, computer memory, and processors
- Output devices, display type (e.g., computer monitor or soft alarm panel), and the data transfer medium (e.g., analog output, HDMI, or wired/wireless network communications)
- The principal design attributes for safety-related applications should be discussed in the context of the architecture: redundancy, isolation devices, and electrical (power) and communications independence.

If the CVML system is intended for mobile applications (e.g., a crash cart or suitcase device) to be utilized during temporary plant conditions, this should be explained. How the CVML system is stored and secured when not in use should be also part of the description.

For every type of CVML model, the literature offers multiple options for addressing the same problem. The software architecture should describe the selected DCNN models, focusing on those aspects that affect the tasks they were designed to perform, and explaining why the selected models were deemed preferable. Given the previously discussed black box nature of the DCNN models, measures similar to the ones described in the previous section should be employed, or otherwise leverage previous efforts' findings and validate them to justify the selected models. For example, in one such effort, one of the better-performing DCNN architectures was composed of five convolutional layers, followed by three fully connected layers running on separate graphics processing units (GPUs) [31], whereas another high-performance DCNN used three convolutional layers with hundreds of feature-extraction filters, followed by two fully connected layers, each with 4,096 connections [36]. The following aspects of the selected DCNN architecture should be described:

- Number of convolutional layers
- Detailed description of the convolutional layers, including the convolution filters at each layer and the stride for each layer (e.g., 5 × 5; stride of 4)
- Pooling functions and number (e.g., average or max pooling), and filter size
- Basis for selecting sizes, stride, and number
- Activation functions (sigmoid, tanh, etc.), types, locations used, and number used
- The number of fully connected layers
- Use of memory-type gates (e.g., LSTM NNs) should be described and justified based on the task the algorithm will be performing
- Output layer (e.g., how many) and data/signal format.

A graphic representing the DCNN architecture should be included to help the reviewer understand the reason for using a single CVML model or ensemble of models. Enough information should be provided to enable the reviewer to determine whether the CVML system is "non-interfering" with other systems that perform safety-related or risk-significant functions (including credited operator actions), and that the hardware/software quality is commensurate with the CVML system's intended application.

2.3.1.1 System Interfaces

Documentation and drawings to illustrate, explain, and justify the data distribution both internal and external to the system should be provided, including all interfaces and hardwired/data communication, whether point-to-point, multiplexed, or networked. This discussion should include those aspects of the design that maintain independence with redundant channels. If a safety-to-non-safety boundary is crossed, the appropriate isolation boundary components should be described including TS, and consequences of failure (or missed failure) on the safety system and the plant.

The input and output interfaces with the plant and plant sensors/actuators, whether hardwired or using some form of data communication should be identified and described including requirements for any required isolation (e.g., between safety-related systems or with non-safety systems/components). The interfaces with control room displays, indicators, controls, and alarm systems should be identified and described, including the system's role and interfaces with post-accident monitoring and any reference by emergency plan-implementing procedures (including credited manual operator actions), as well as requirements for any required isolation. For the CVML system, input interfaces should address the required camera inputs (e.g., still or moving image). Interfaces with other plant systems that provide input (e.g., fire detection/suppression systems) or data communication avenues via network connections should be described, including wired or wireless network communications (e.g., Ethernet), HDMI cables from the computer to the local display, and analog outputs to alarm relays or annunciator panels.

Additionally, the HSIs for the licensee's user, maintenance, and engineering workstations used for test and maintenance should be identified and described, whether considered internal or external to the new plant system. This includes requirements for any required isolation, as well as addressing changes from the existing HSIs. The human factors engineering processes and results used for compliance with IEEE Std 603-1991 should be referenced or described. Procedures should be established to control actions (such as those taken by the fire brigade after being notified of a fire by the control room). The procedures should encompass actions to be taken by plant personnel upon report of an alarm or event, as well as actions to be taken when the CVML system triggers a diagnostic alarm, indicating that operability may be compromised. The CVML system interface should also account for the possibility of indicating false positives on the interfaced systems, in addition to methods of mitigating them.

All safety-to-non-safety interfaces should have adequate electrical isolation based on the electrical hazards present and the electrical isolation criteria in Standard Review Plan (SRP) BTP 7-11, "Guidance

on Application and Qualification of Isolation Devices." The use of power sources, both electric and non-electric, should be described and should state how the system continues to perform its credited safety functions while power sources are bypassed for maintenance to comply with IEEE Std 603-1991. The primary and backup electrical power sources should be identified, such as a site 120VAC (e.g., wall power) source and battery backup. If a safety-related electrical power source is required, appropriate provisions for backup power with safety-to-non-safety isolation should be described, if applicable. The consequences of failure of the electrical power sources should be described and mitigated for safety-related applications.

2.3.1.2 Redundancy

Redundancy helps ensure that a single failure will not impair a safety system's ability to perform safety functions, as per the GDC of Appendix A, "General Design Criteria for Nuclear Power Plants," to 10 CFR Part 50. Redundancy means that no single failure can prevent the CVML system from performing its task. For safety-related and certain risk-significant applications, this means that sufficient redundancy should be afforded from input to the output devices, such that the CVML system can perform its design function despite a single failure, even single failures in the processor executing the CVML model.

Failure analysis of the CVML system should be performed to identify all single failures that can prevent the CVML system from performing the task for which it was designed. Such failure analysis should be performed on each major component of the system, providing enough detail to identify and mitigate single failures. The system architecture should conform to the guidance in RG 1.53, "Application of the Single-Failure Criterion to Nuclear Power Plant Protection Systems" [42], which endorses IEEE Std 379, "IEEE Standard for Application of the Single-Failure Criterion to Nuclear Power Generating Station Safety Systems" [43]. The critical issue is that the redundant architecture will perform its safety function(s) in the event of a single failure and all undetectable failures.

A failure modes and effects analysis (FMEA) should be implemented to demonstrate that the use and application of redundancy in the new architecture ensures that the safety functions can be achieved in the event of a postulated single failure in the presence of all undetectable failures[h]. The FMEA should document the postulated failures and their effects on both the plant and the system. It should also provide sufficient detail when evaluating failures such as failure of an individual input, input module, processing module, output module, individual output, or any communications portion of the system. The new architecture should meet the criteria of IEEE Std 603-1991 when applying the associated guidance of IEEE Std 7-4.3.2-2003.

From an input perspective, redundant image sensors (cameras) or electrical signals provided to connected equipment may be necessary to ensure that loss of video feed does not impair the functioning of the CVML system. Postulated failures of the CPU that executes the CVML model would require that appropriate compensatory measures be identified. The means by which catastrophic failure of the processor is detected, alarmed, and mitigated should be discussed for safety-related and risk-significant applications of the CVML system. CPU outputs that are provided to operators (e.g., operator displays or alarm panels) or electrical signals provided to connected equipment relied upon to prevent or mitigate accidents will also require sufficient redundancy to ensure that the safety function(s) can be performed. The design should also include sufficient redundancy to enable system testing. Redundancy in the CVML system should also allow for performing any maintenance on the system if the system will be deployed for periods exceeding the proof-test interval of any of its critical components.

A redundant CVML model may be cost prohibitive, in which case the operator(s) may need to resume performance of the task. For applications that do feature non-trivial consequences (e.g., measurable cost

---

[h] An "undetectable failure" is neither observable by an operator (e.g., front-panel indications such as displays and status/alarm lights) nor captured by the system's built-in diagnostics or automated self-test features. An undetectable failure may be a latent failure due to inadequate diagnostic test coverage, and can defeat redundancy features.

impact), such as an SIL-2-rated application not requiring redundancy, some form of alarm or operator notification is needed to prompt the operator to resume the activity that had previously been performed by the CVML system. For example, immediate redundancy might not be required for applications that entail once-a-day monitoring of a gauge via AGRS, and operators could compensate for it as part of their rounds as long as the AGRS failure is diagnosed and reported.

2.3.1.3 Independence

Independence is a design characteristic of an architecture that prevents failures from propagating (1) from one safety-related system to another, (2) between redundant portions of a single safety-related system, and (3) from a non-safety-related system to a safety-related system. Sufficient independence should be incorporated to ensure the effectiveness of the redundancy and diversity in the DI&C to maximize the reliability of systems that support safety-significant functions. Independence complies with GDC of Appendix A, "General Design Criteria for Nuclear Power Plants," to 10 CFR Part 50. The new architecture should demonstrate the relevant clauses of IEEE Std 603-1991 when applying the associated guidance of IEEE Std 7-4.3.2-2003. The architecture should conform to the guidance in RG 1.75, "Criteria for Independence of Electrical Safety Systems," which endorses IEEE Std 384, "IEEE Standard Criteria for Independence of Class 1E Equipment and Circuits." Additional guidance regarding communications independence is provided by the staff in Interim Staff Guidance (ISG) document DI&C-ISG-04, "Highly Integrated Control Rooms and Digital Communication Systems."

The three recognized types of independence: physical, electrical, and communications should be addressed. Physical independence is physical separation or barriers between independent systems, such that failures do not propagate and prevent performance of the credited safety functions. The CVML system should have physical barriers that protect it from accident events occurring in the location where it is installed. For example, a seismic event could cause mechanical failure of a nearby system, potentially impacting the CVML system. Another example is a water spray, steam release, or oil spill that could impair the CVML input sensors, preventing the CVML system from performing its design function. To guard against these events, the CVML system should have sufficient physical barriers to ensure that the design task can be performed.

Electrical independence is the electrical isolation between the independent systems, such that an electrical fault originating from one safety-related system—or equipment that is not safety-related—will not adversely impact a safety function. The electrical isolation devices or measures installed to prevent electrical fault propagation should be qualified as part of the safety-related system. The CVML system should be isolated against electrical power faults originating from outside the system. The electrical isolation should also ensure that a fault does not propagate between redundant power supply circuits in the system, such as between the primary power source (e.g., 120VAC) and the backup power source (VDC). For mobile systems (e.g., crash cart or suitcase device), this may not apply, as they will be powered by a temporary source (e.g., battery powered).

Communications independence entails separation between the independent systems such that communications failures originating from outside the safety system will not adversely impact the safety function. The new architecture should ensure that physically and electrically independent portions of safety systems do not depend on information from other independent portions of the safety system, or from outside the safety system (functional independence). The safety-related system should be unaffected by potential failures in the communications mechanisms (devices, physical interconnections, and protocols) or in the information being communicated. Adequate controls should be in place to prevent these failures from propagating to the safety-related system. The information received from outside the safety division, spurious actuations of I&C equipment due to credible failures, or consequential actions of non-safety systems should not adversely impact the safety function. The CVML system should offer isolation against failures in the communications portion of the system (e.g., physical cables, network devices, and communications protocols). If any of the information originates outside the CVML system,

the system should not depend on that data to perform its design task (i.e., not the input sensor data). If the CVML model (DCNN) can be corrupted by data when in online learning mode, that mode should be disabled (i.e., batch learning only). If operator responses are allowed by the CVML system, then slow or erroneous operator input should be considered in the failure analysis and appropriately protected against to satisfy the requirement for communications independence.

A unique aspect of CVML models that was discussed in Section 2.1 and 2.2 is their use of common datasets and feature extraction engines, potentially compromising the independence of the CVML models and introducing forms of common cause failure (CCF). This should be addressed as discussed in Section 2.1 and 2.2.

2.3.1.4 Deterministic Behavior

Deterministic behavior ensures the predictable, repeatable behavior of systems that perform safety functions, as per the GDC of Appendix A, "General Design Criteria for Nuclear Power Plants," to 10 CFR Part 50. The new architecture should meet the relevant clauses of IEEE Std 603-1991 and the associated guidance of IEEE Std 7-4.3.2-2003.

The new architecture input signals and system characteristics should result in output signals through known relationships among system states and responses thereto. It can be argued that CVML models are known relationships among system states and responses thereto, but they are incomprehensible to humans and are stochastic to some extent.

To enable timely completion of credited actions, the system should produce the same outputs for a given set of input signals (and the sequence of inputs) within well-defined response time limits. Assurance that the same input(s) result in the same output(s) is challenging to demonstrate. How "same" is defined and measured could be subjective, given potential variation in plant conditions and the high-level reasoning capability of the CVML model. Though certain portions of the input-process-output loop of the CVML system operate in predictable ways and provide the same output for the same set of inputs (e.g., the camera input sensors), the high-level reasoning of the CVML model makes the system inherently non-deterministic. Nonetheless, the predesigned test cases with acceptance criteria for documenting the CVML system's performance could be leveraged as a benchmark for future potential modifications to the system. The results of acceptance testing will allow the reviewer to assess this aspect of determinism and perhaps establish an acceptance threshold.

The design should adequately identify and account for hazards that could challenge the predicted behavior. Any hazards that impact CVML system behavior should be determined (e.g., in the context of the measures above), and identify any design attributes, maintenance activities, or operator actions that mitigate their consequence(s). This includes the worst-case response times to a decision output for a given change in the input. Some plant conditions may delay the CVML system time to decision, producing adverse consequences (e.g., too long to recognize a fire, or failure to recognize that a process parameter exceeded its limit in an analog gauge reading). The deterministic behavior of digital data communication outputs should ensure that system timing requirements derived from DBE analyses are satisfied by the replacement system architecture. It should also ensure that the replacement system architecture and communication protocols provide features to guarantee that the system generates the correct response to inputs within the time credited to produce a response. A worst-case time-to-decision measure should be established to enable operators to determine when human intervention is necessary to complete the design task. The worst-case time to decision would be another benchmark against which to compare future changes.

2.3.1.5 Diversity and Defense-in-Depth

D3 are methods of protecting against CCFs, as per the GDC of Appendix A, "General Design Criteria for Nuclear Power Plants," to 10 CFR Part 50. The D3 assessment should conform to the guidance in SRP BTP 7-19, "Guidance for Evaluation of Diversity and Defense-in-Depth in Digital Computer-Based

Instrumentation and Control Systems," including use of an analysis as described in NUREG/CR-6303, "Method for Performing Diversity and Defense-in-Depth Analyses of Reactor Protection Systems." When a facility change is being performed under 10 CFR Part 50.59, the guidance in Regulatory Information Summary RIS-2002-22, Supplement 1 can be used in the CCF analysis. If the CCF analysis determines the probability of CCF to be sufficiently low, it can be removed from consideration.

The relevant CCFs should be identified, and postulated accident events concurrent with the CCF should be adequately addressed via diverse methods (including operator action within the credited response time), and the diverse methods used should be adequate (e.g., tool, human, and signal diversity). While diversity is an aspect of the fundamental I&C design principles, it is considered only one potential means of addressing CCFs. The diversity should be discussed in terms of how it supports the defense-in-depth assessment and other measures to address CCFs. The design should address potential CCFs due to:

1. Systematic faults caused by design and implementation defects within redundant divisions of safety-related systems; for example, a CVML system with diverse and redundant portions that nonetheless utilizes the same software and could simultaneously fail in the same manner, thus preventing the safety function
2. Faults that are propagated from non-safety systems to safety-related systems, and that can adversely impact the safety-related systems
3. Internal and external hazards that can adversely impact a safety-related system, or systems belonging to multiple levels of defense.

Faults in the second category can be addressed via independence among safety systems as well as independence between safety-related and non-safety systems. Faults in the third category can be addressed by an effective qualification program. Faults in the first category are more challenging. Systematic faults can be addressed via diversity and good design practices and provisions (e.g., physical separation). For example, the fire watch use case would implement diverse input cameras (e.g., one visible light camera and one infrared camera, or two different brands of cameras). However, this is more challenging when considering CVML models (as opposed to a typical DI&C). One common concept in CVML is software reuse, whereby software code (or portions of code from a previously commissioned and operating digital system) is applied to a new but similar software project. The concept of software reuse as applied to a CVML model is considered transfer learning, whereby an existing model is applied to a different problem. Such sharing is one of the many advantages of digital systems, but it raises a key concern: a design using shared data or code can potentially propagate a CCF or common mode failure via software errors, defeating the redundancy achieved by the hardware architecture. This sharing, in turn, increases the consequences of failure of a single module and reduces the amount of diversity available within a single safety channel. One solution is to use different CVML models with different architectures (and datasets), the outputs of which would be voted/compared prior to operator notification or alarm. However, given the relative complexity of CVML models, as well as the overlapping, underlying concept of using DCNNs and overlapping data often available in public repositories, establishing complete independence to ensure diversity might prove harder to demonstrate. Alternatively, "diverse function" could take the form of a completely different system, such as a thermal camera for a fire watch CVML system, or the operator resuming performance of the task upon failure of the AGRS system. This could also serve as a defense-in-depth strategy when multiple diverse systems are deployed.

Other systematic faults must also be considered. For example, plant configurations that would require deployment of multiple CVML systems (e.g., using multiple crash carts or suitcase devices), in addition to multiple CCF sources, must be considered, particularly for electromagnetic and radio frequency interference upsets that could impact multiple staging locations of the portable systems.

2.3.1.6 Simplicity of Design

The simplicity of the design will facilitate efficient engineering analysis of the safety of the I&C design. For example, complex DI&C systems can present a challenge in terms of demonstrating conformance to safety-related system design criteria such as independence. In this context, simple design concepts support straightforward engineering analysis and testing of DI&C systems to ensure that defense-in-depth measures were appropriately implemented. Reduced design complexity facilitates the review and enables the reviewer to reach a conclusion faster and with less ambivalence. The relevant criteria found in IEEE Std 603-1991 should be addressed.

Many decisions are made during the design of a system. Some can lead to added complexity in the final configuration. When considering the inherent complexity of the CVML models themselves, and the various attempts to reduce the training time and improve task performance [31, 34, 35, 38], the CVML systems can be quite complex. The review may necessarily require analyzing the CVML model (i.e., getting inside the black box) to provide a detailed description of the internal image processing (e.g., performed by the convolutional layers) and high-level decision making of LSTM NNs (hidden layers), giving a better understanding of, among other things, numerical errors (and their contribution to bias) and potential failure modes.

Evaluation of the simplicity-of-design principle relies more heavily on engineering judgment than do other fundamental design principles. Given the subjective nature of this principle and the reliance on engineering judgment, any aspects of the design that satisfy this consideration should be highlighted. The process used to minimize complexity and reduce the CVML model's overall footprint (e.g., in terms of required memory) should be highlighted.

For design decisions resulting in more complex approaches than might otherwise have been selected, the benefit(s) obtained should justify the added complexity, particularly with respect to the other fundamental design principles. It should consider providing a rationale for any design decision that resulted in the replacement system architecture being more complex. In this context, the added complexity should (1) provide a safety benefit and (2) not diminish the design's conformance to the fundamental I&C design principles. Design decisions involving such tradeoffs may be driven by the need to satisfy a safety requirement (e.g., surveillance testing or improved maintainability or operability during faulted conditions).

2.3.2 System Functions

The functions of the new system should be developed and described. Each design function's description should include equipment from sensor to actuator or display device(s), if applicable, including logical operation, manual versus automatic, and any interdependencies (e.g., signal split and use in a safety function as well as in a control room display). These design functions are safety functions implemented in the application-specific software, programmable logic, hardware (e.g., hardware voters and relays in traditional I&C safety systems), credited manual operator actions, or some combination thereof. Adequately addressing this ensures compliance with the relevant criteria in IEEE Std 603-1991 and the associated guidance of IEEE Std 7-4.3.2-2003.

The CVML system function description should identify the functions performed and their safety or risk significance (e.g., the specific function of AGRS at a certain location, the measurement being acquired, its safety-related role, and the CVML system's impact on plant safety). The safety classification of each safety function should be described, along with whether other functions cause independence constraints based on the safety classifications. Any associated trip/actuation functions credited for each anticipated operational occurrence and postulated accident should be identified, as well as input/output ranges and setpoints (including defining all applicable uncertainties in the complete loop).

All monitored variables (i.e., those pertinent to controlling each protective action or for credited operator actions) should be identified and described along with methods of input processing (e.g., for

CVML, the information on the image/video format of the image collection process should be provided). This should include size and memory requirements, so reviewers can confirm that the hardware and software are compatible. The minimum number and location of sensors and equipment being relied on for the safety function (i.e., pertinent to the monitored protection functions or for credited operator actions) should be identified and described. For the CVML system, the number, placement, and location of the cameras should be described.

Operational performance, including accuracy and response times (where appropriate, performance requirements are defined for different initial plant conditions and any relevant DBEs) of the CVML system should be described, as well. Specifically, the input processing time (e.g., frame rate [fps] and limitations on image size), time necessary to make a decision (i.e., time to process the input data and generate the decision at the output layer), and, if significant, output time (when the output is transmitted over a network to a remote operator display, the control room, or remote alarm panel representing a worst-case time to decision) should be discussed. The rate of false positives and false negatives, how they can be mitigated to reduce the probability of spurious actions (e.g., the appropriate level of signal filtering/validation to minimize the potential for spurious actions based on the monitored variable), and how to ensure there are no missed actions should be discussed.

The range of transient and steady-state conditions throughout which the safety-related application should perform should be described, including conditions (e.g., plant process) that can potentially degrade the performance of the safety function performance. This includes environmental and operational limitations related to, among other things, the cameras. For example, camera sensor inputs may have minimum lighting requirements or a maximum detectable frequency and related issues—such as a vibrating pointer on analog gauges, or the maximum distances to the objects being detected.

2.3.2.1 Functional Allocation

The functions (e.g., logic functions) distributed within physical hardware should be described using text and drawings. The mapping of logic drawings (i.e., functions) to logic elements in the system should be described. The mapping of design functions and auxiliary features to software, hardware, manual operating actions, or some combination thereof should be demonstrated. System architecture drawings could be used to graphically present the functions allocated to the hardware and software. For third-party components (e.g., camera sensors), high-level descriptions are sufficient—that is, image sensing and data formatting is acceptable. However, for internal functions essential for the CVML system to perform its design task, detailed descriptions would be required. These should detail the model's internal architecture (e.g., number of layers, size of each layer, activation functions), the internal function of each part of the model, and the hardware-software interactions. The relevant criteria in IEEE Std 603-1991 and the associated guidance of IEEE Std 7-4.3.2-2003 should be addressed.

The allocation of design and service/test functions to the various elements of the proposed architecture (e.g., hardware, software, and operators using HSIs) should be demonstrated. This includes other software critical to the proper functioning of the CVML system, such as embedded software for camera control, digital signal processing algorithms for noise filtering, and custom software for the display function (e.g., an operator display). The other functions assigned to the software, including internal diagnostic, self-test, and maintenance functions (using text and graphics) should be discussed. The design's response times and how they fall within the range credited in analyses of the applicable modes of operation should be demonstrated when considering conditions in which response times are critical to the safety function. Guidance can be found in SRP BTP 7-21, "Guidance on Digital Computer Real-Time Performance," of NUREG-0800.

User-allocated functions (e.g., calibration, alarm reset, bypass feature, or data retrieval) should be described. Any user action credited in the NPP safety analysis—or that is critical to the continued functioning of the CVML system—should be appropriately highlighted, and worst-case response times should be specified and justified (e.g., using calculations, as appropriate). At minimum, the CVML

system should maintain a response-time performance equivalent to that of a human operator. If the CVML system impacts credited operator actions, the NPP FSAR would need to be updated and documented.

2.3.3 Fault Detection/Diagnostics

The test functions should be detailed and described as well as the design functions. Test functions are digital equipment features that support the design functions. The test functions, unlike the design functions, do not directly relate to the performance of safety functions, but rather the specific activities of the digital equipment, including the functions necessary for operation, periodic testing, self-test, self-diagnostics, and maintaining a secure operational environment. For each test function, the relevant criteria in IEEE Std 603-1991 and the associated guidance of IEEE Std 7-4.3.2-2003 should be addressed. Additional guidance on self-test features is provided in SRP BTP 7-17, "Guidance on Self-Test and Surveillance Test Provisions."

The discussion of self-tests and self-diagnostics should demonstrate compliance to any proposed TS for new system functions, if applicable. The means to detect malfunctions should be demonstrated. The following are examples of issues that may apply to the replacement design:

How the application interfaces with and uses the self-test and self-diagnostic features. This includes the quality of the decisions (i.e., high-level reasoning) of the CVML system. For example, corruption of the CVML model or image can occur, either due to corruption of the executing program or a hardware glitch that causes the CVML model to initiate a calculation error that can significantly impact its performance. A discussion of the malfunction detection coverage should be included, considering the automated features in combination with TS surveillances (i.e., periodic functional checks of the CVML system for undetected failures, such as frozen frames, while deployed in the field), if applicable. This includes assessing the operability of the CVML system in the wake of certain natural phenomena and potential man-made events, as well as conditions that require the CVML system to be declared inoperable and for certain actions to be taken, such as reverting to a human operator. The CVML system should provide a means of verifying its proper operation without removing it from service. Additional actions necessary to ensure that critical functions are within the test coverage envelope need to be defined along with any relevant proof-test intervals. For example, CVML system power supplies may have a minimum test interval to ensure that the reliability calculations are valid. The approach to periodic testing and the test intervals should be identified and specified. The test methodology should be described. While this may be achievable in the case of AGRS (by means of fabricated gauges or manually altered gauges for testing), it may be more challenging for use cases that are harder to create, such as starting fires. Instead, videos of fires can be positioned within the camera's perspective, but this imposes some image quality challenges that might impact the testing process. It is also possible to test the various parts of the system individually (e.g., test the CVML using digital videos fed directly to the model); however, a use case must be developed on how the individual testing of the system component provides sufficient detail and coverage to replace the full system test.

How the design implements communication messages, if applicable. The communication method should not affect interfacing systems, nor cause the CVML system to fail to perform its intended functions. As such, any communications-related hazards should be identified and appropriately mitigated. This should include a description of the types and purposes of each message, the format of each message, the response of the receiver to invalid data, methods of detecting repeated messages, alarming on malfunctions, and the application of each message (e.g., voting or bypass). Error detection means provided in communication messaging and processing apply to communication messages used by safety functions. This includes a communication check (e.g., a watchdog timer implementation) to ensure that the data received are live, as well as an error-detecting method to ensure the data are valid. For each deployment, credible failure modes of the communications method should not interfere with connected equipment, either through corrupted wired communications or as a result of electromagnetic and radio

frequency interference upsets via wireless communications, and that communication pathways—particularly wireless—are protected from both unintended and malicious faults. Conformance to the DI&C-ISG-04 assessment containing these data should be demonstrated.

How the design prevents software failures from affecting the watchdog timer timing[i] *and timeout, if applicable*. This should address hardware and software malfunction coverage for the watchdog timers, including a description of the annunciation and the effects on the plant during and after any reset function initiated by an expiring watchdog timer.

Treatment and detection of malfunctions in the system. This involves inputs (e.g., data from sensors and transmitters) to the system logic, including internal voters (e.g., triple-modular redundant architecture), if applicable, within the safety-related application, and malfunctions in the system outputs. This should include the expected failure state(s) of each input and loss of input(s), and image quality. This may also include data corruption after the input(s) are processed and formatted for use by the CVML algorithm (e.g., RGB values of the image data that were corrupted prior to the CNN's convolution operation). Consideration should be given to the system response to each failure, the expected failure state(s) of each output, and the response of the plant and operators to each failure. This includes "delayed" or "no decision" outcomes by the CVML model, perhaps due to glitches or errors in the functioning of the NN's hidden layers. Critical parameters for the functioning of the CVML system should be identified, such as CPU temperature and power supply voltage, and be based on the application FMEA. The malfunctions should be detected and mitigated, as well as any other potential failures identified during the failure analysis.

2.3.4 Maintainability

The service functions along with the design functions should be defined. Service functions are digital equipment features that support the design functions. The service functions, unlike the design functions, do not directly relate to the performance of safety functions, but rather to the specific activities of the digital equipment, including the functions necessary for configurating, validating, maintaining, and incorporating design modifications. The safety function should be protected, both during maintenance and in the presence of a single fault. In both use cases featured in this report, this is accomplished by the operator taking over the CVML system role while the system undergoes maintenance.

Maintenance/service functions should be built into the CVML system, and any special test tools, such as software tools, should be provided to NPP staff for system maintenance. This could involve migrating the CVML system to a new camera via installation of a CVML software module upgrade. Service functions may require a separate maintenance interface, which can be either temporary or permanent.

Hazards associated with the maintenance functions should be addressed (e.g., incorrect training inputs) and mitigating measures identified. The maintenance/service function should include the capability to verify the proper operation of the CVML model and retraining of the model if plant conditions require it. This could include, among other things, using a test input to confirm the proper functioning of the system after maintenance is conducted. For an image classification system (e.g., fire detection), a pre-defined image can be used as calibration input, with the actual output verified against the expected output.

[i] A watchdog timer is a special function that monitors the execution of a computer system to guard against system freeze, resource lock-up, or slow execution (as compared to a specified maximum time interval). As the name implies, the watchdog function counts based on input from the system RTC, and when a threshold is reached (i.e., the timer times out), the watchdog resets the internal CPU to reboot the system, free resources, etc. A watchdog timer can be executed in software, but the most reliable implementations are hardware components external to the CPU so as not to be impacted by a common mode failure of software.

2.4 Software Quality

The typical acceptance of software for safety system functions is based on (1) confirmation that acceptable plans were prepared to control software development activities, (2) evidence that the plans were followed throughout an acceptable software lifecycle, and (3) evidence that the process produced acceptable design outputs. It is assumed that the software development plans are implemented within a QA program that conforms to safety requirements. To this end, the developer of the CVML system should have a defined QA program that conforms to the requirements of 10 CFR Part 50, Appendix B. Guidance on software QA can be found in IEEE Std 7-4.3.2, "IEEE Standard Criteria for Digital Computers in Safety Systems of Nuclear Power Generating Stations," endorsed by RG 1.152. Guidance on software quality management can be found in IEEE Std 1074, "IEEE Standard for Developing Software Lifecycle Processes," endorsed by RG 1.173.

The proposed framework being used to design and develop the application should be described. This framework should supplement the overall QA program descriptions with specific system, hardware, and software development activities, including a description of the proposed development lifecycles, development documents to be produced, and management activities to be implemented in the design and development of applications. The framework should describe the following system development process activities:

1. Create the concepts on which the system design will be based
2. Translate these concepts into system requirements
3. Allocate system requirements to system elements (e.g., software, hardware, and HSIs)
4. Implement the design into hardware and software functions
5. Integrate system elements (e.g., software and hardware)
6. Test the unit functions and the completed system to confirm that system requirements were implemented correctly
7. Analyze hazards and incorporate requirements that eliminate or mitigate identified hazards throughout the development process
8. Perform validation and verification (V&V) activities on work products throughout the development process.

The software lifecycle process should follow the guidance in RG 1.173, "Developing Software Life Cycle Processes for Digital Computer Software Used in Safety Systems of Nuclear Power Plants," [44] or alternately, describe how the requirements referenced in RG 1.173 are satisfied. Demonstration of the software lifecycle process includes describing the software lifecycle processes to be used, identifying any planned exceptions and clarifications, and describing how these planned exceptions and clarifications meet the underlying requirements. Additionally, there may be a situation in which a commercial-grade item that utilizes software will be dedicated for use in a safety-related application (e.g., using a preexisting feature extraction engine or object recognition tool such as You Only Look Once). The software being dedicated for safety-related use by a vendor must demonstrate that the development lifecycle for the commercial-grade software meets the above software quality requirements. If necessary, the design documentation and test outputs for the commercial-grade software will have to be reconstituted as part of the dedication process. The vendor oversight plan in order to ensure that the vendor executes the project in a manner consistent with the requirements should be described and the American Society of Mechanical Engineers (ASME) standard Nuclear Quality Assurance (NQA)-1, Part II, "Quality Assurance Requirements for Nuclear Facility Applications," Subpart 2.7, "Quality Assurance Requirements for Computer Software for Nuclear Facility Applications" [45]. This ASME standard is also endorsed by the NRC staff. Subpart 2.7 is specific to digital technology and software.

Typically, the software organization will develop a software QA plan (SQAP) for implementation under an approved QA program, such as one conforming to RG 1.28 [46], which endorses ASME NQA-1, and the ASME NQA-1a Addenda, "Addenda to ASME NQA-1-2008 Quality Assurance Requirements for Nuclear Facility Applications." The SQAP should demonstrate compliance with the requirements of 10 CFR Part 50, Appendix B and the overall QA program. The SQAP would typically:

- Identify which QA procedures are applicable to specific software processes
- Identify the particular methods chosen to implement the QA procedural requirements
- Augment and supplement the software QA program, as needed.

Additional Guidance on SQAPs can be found in NUREG/CR-6101, "Software Reliability and Safety in Nuclear Reactor Protection Systems" [47].

The structure of the software QA organization should be described and should guarantee sufficient authority and organizational freedom—including sufficient independence from cost and scheduling —to ensure that the QA organization's effectiveness is not compromised.

If software QA requirements are not adopted in R&D of the CVML model, since R&D is an iterative process that is often less systematic than a software product, the CVML software development may need to be restarted following R&D software development (without using the R&D-generated software but leveraging the R&D knowledge gained) to ensure that the end software is developed in accordance with a QA program implemented as part of an approved Appendix B program.

2.4.1 Development Lifecycle

The development of safety system software should progress according to a formally defined lifecycle. Many lifecycles have been defined in the technical literature and in national and international standards. All software development lifecycles share certain characteristics, and the performed activities are common to all lifecycles, but they differ in their definitions of lifecycle activities and the order in which such activities are performed. Lifecycle activities produce process documents and design outputs that can be reviewed and assessed. The software lifecycle should be selected and documented and should specify those products that will be produced over the course of that lifecycle. Further information on lifecycles and process document contents is found in NUREG/CR-6101.

The project management or organizational processes that will be employed by the QA program should be described and used to define the project's organization, planning, execution, monitoring, control, and closure of the development effort. The project management function should incorporate the impacts of changes to the project risks, schedule, and budget. The project personnel should be trained on the lifecycle model processes and procedures, with periodic management oversight reviews (e.g., stage gates) to ensure adequate progress and conformance to organizational/quality goals and design objectives. The software management process should be described, using information provided in SRP BTP 7-14 as a guide. The description of the organizational and project management processes should include the following:

- Measures for creating plans to control the system development environment, including hardware and software in accordance with 10 CFR Part 50, Appendix B, Criterion V, "Instructions, Procedures, and Drawings," with the planning process resulting in a set of documents to be used for controlling and overseeing the development of system elements, including hardware and software
- Controls for identifying the project scope, deliverables, lines of communication, formal and informal reviews, and interfaces with other internal and external organizations
- Provisions for the establishment, documentation, and maintenance of a schedule that considers the overall project, as well as milestones

- Provisions for risk management, including problem identification, impact assessment, and development of mitigation plans for risks that could significantly impact system quality goals, along with appropriate metrics for tracking resolution progress, as well as the additional guidance on software-related project risk activities contained in IEEE Std 7-4.3.2-2003
- Establishment of quality metrics throughout the software development lifecycle in order to assess whether the quality requirements of IEEE Std 603-1991 are being met, with additional guidance from IEEE Std 7-4.3.2-2003
- Adequate control of software tools to support system development and software V&V processes, with additional guidance contained in IEEE Std 7-4.3.2-2003
- Provisions for documenting and resolving problems and nonconformances in the system elements
- Provisions for effectively overseeing lifecycle activities.

In every aspect of design, all changes and modifications should be documented and controlled in accordance with the quality program, as per 10 CFR Part 50, Appendix B, Criterion III. The project management function should incorporate the impacts of changes to the project risks, schedule, and budget.

2.4.2 Software Verification and Validation

V&V refers to the set of activities meant to confirm that design outputs meet the technical requirements specified in the functional and requirements specifications documents (e.g., system, hardware, and software, as appropriate). The V&V process involves independent reviews of software outputs, including the tracing of requirements between the software design input and output documents. V&V also tests the software components and integrated hardware/software system by using independently developed test procedures and the test metrics introduced in the design verification section. Acceptance criteria can be defined by the designer, but the V&V team independently confirms the acceptance thresholds have been satisfied. The V&V effort should be described and show that it is sufficiently disciplined and rigorous to ensure a high-quality software development process. The software V&V processes should address V&V organization responsibilities, processes, activities and tasks, reporting, administrative controls for anomaly resolution and reporting, task iteration policy, deviation policy, and test documentation. The applicable review guidance for software V&V processes is found in SRP BTP 7-14, which references IEEE Std 1012, as endorsed by RG 1.168.

It is acceptable to adapt software V&V activities and tasks to reflect important process differences, technology differences, and exceptions related to the use of integrated design environment tools. As discussed, the CVML system's lack of repeatability (e.g., unreproducible, unpredictable results) from one CVML model or developer to the next, due to the stochastics nature (see Section 2.3.1.4) of the CVML learning process, is a key consideration during the V&V process. This issue stems from the underlying ML development and training processes, resulting in a given CVML model potentially outputting different results for the same input. This emphasizes the importance of establishing well-defined acceptance criteria during the (system and software) V&V process, such that repeatability is ensured when implemented in the NPP target environment.

The V&V organization's level of independence should be demonstrated, as it is an essential aspect of the software QA program in accordance with 10 CFR Part 50, Appendix B, Criterion I. The V&V team should be granted managerial, technical, and financial independence (per RG 1.168). V&V personnel should not be subject to scheduling constraints or pressure from the software designers or project managers for reports or review efforts, and the V&V team should report to a level of management that does not exert direct pressure for a favorable V&V report. This does not require V&V to be performed by a separate company, but for the V&V staff to have sufficient independence to raise concerns, regardless of schedule impact and without fear of retaliation, ensuring that sufficient financial resources are

available to perform an adequate level of design verification in accordance with 10 CFR Part 50, Appendix B.

As discussed in Section 2.1 and 2.2, establishing independence in CVML software V&V is challenging (software V&V falls under design verification as defined in Criterion III of Appendix B). The software would likely rely on overlapping concepts and datasets, and techniques must be developed to address this concern. On the data front, methods such as the generative adversarial network could be used to generate datasets that are guaranteed unique. Alternatively, data augmentation through regularization increases the size of the training dataset in order to improve the training of the model [31, 34, 38]. These techniques manipulate the dataset (e.g., by rotating an object in an image frame) to present diverse data to the model during training in order to improve task performance and minimize the chances of overfitting. This is a helpful technique when training data are scarce. For applications aimed at NPPs, use of data augmentation should be documented and the bases for acceptability addressed. As for the models, validation could rely on different packages or tools (e.g., using PyTorch versus using TensorFlow), in addition to using different environments, etc.

The number and quality of the V&V personnel should be addressed. Staff performing V&V activities should at least be as knowledgeable as the staff designing the CVML software. There is no specific requirement regarding the number of V&V personnel, but the industry rule of thumb is that a good V&V process requires as much work as the original design effort. Also, V&V engineers should be qualified to understand the process, technology, and software.

The V&V effort should be fully and carefully documented to ensure that all discrepancies are documented in a report, and that those discrepancies have been resolved, with the resolution documented, tested, and accepted by the V&V organization. A significant contributor to issues that arise in final products is problem "fixes" that did not go through the V&V process and were not properly tested, resulting in additional problems created by the fix not being found.

The V&V reporting requirements should be described. It should highlight that output reports document all V&V activities, including the personnel conducting the activities, the procedures, and the results. This includes review documentation requirements, evaluation criteria, error reporting, and anomaly resolution procedures. V&V reports should summarize the positive practices and findings, as well as the negative ones. The reports should summarize the actions performed and the methods and tools used.

2.4.3 Software Configuration Management

Software CM is another important process, because software errors can occur when software changes are made to the wrong version of the software, or when the changes are not sufficiently tested to ensure they do not introduce new errors. CM begins once the initial software is first released by the software design group, and extends to the testing and verification of modified software or documentation, and to identifying the organization responsible for performing the testing and verification. As such, the CM processes should be described, including how they follow the guidance in RG 1.169, "Configuration Management Plans for Digital Computer Software Used in Safety Systems of Nuclear Power Plants" [48], or alternately, describe how the requirements referenced in RG 1.169 are satisfied. The applicable criteria for CM processes can be found in IEEE Std 828, as endorsed by RG 1.169, "Configuration Management Plans for Digital Computer Software Used in Safety Systems of Nuclear Power Plants." SRP BTP 7-14, Sections B.3.1.11.4 and B.3.2.3, reference both the IEEE standard and RG 1.169. RG 1.152 endorses IEEE Std 7-4.3.2-2003, subject to the positions and modifications identified in the RG. IEEE Std 7-4.3.2-2003, Clause 5.3.5, provides guidance on CM.

A critical item to look for is an exact definition of who will control the software (e.g., the group responsible for maintaining the various versions of the software, giving out the current version for testing or modification, and receiving back the modified and tested software). The organizational responsibilities

for CM and for any designated configuration control board should be described. The responsibilities and authorities for managing and accomplishing the planned CM activities should also be described, applying CM to the project, coordinating CM activities with the other project activities, and tools used for CM activities.

All software—not just the operational code to be used in the safety application—should be controlled. This includes any software or software information that affects the safety-related software, including:

- Software requirements, designs, and code
- Support software used in development
- Libraries of software components essential to safety
- Software plans that could affect quality
- Test software requirements, designs, or code used in testing
- Test results used to qualify software
- Analyses and results used to qualify software
- Software documentation
- Databases and software configuration data
- Pre-developed software items that function as safety-related system software
- Software change documentation
- Tools used in the software project for management, development, or QA tasks.

CM is especially challenging when using publicly available models that introduce revisions on periodic bases. As discussed earlier, for CVML models, one important consideration related to CM (including version control) is the publicly available databases and models [34, 38] that may be used during development. The publicly sourced software components and test/validation data utilized in the development should be placed under configuration control, whereby the initial version is recorded and protected from inadvertent, incorrect, and unauthorized changes. This ensures traceability to a known source, as well as tracking of all changes made thereafter during the development process.

Additionally, the model or algorithm itself may not require modification (e.g., when sourced from a known public repository), but the support software functions (e.g., maintenance features, diagnostic functions, or self-test features) may undergo changes throughout the development lifecycle. This is critical to requirements management and efficient project execution (i.e., less rework), as this ensures that, as the design proceeds to completion, the final design and implementation will support the plant's needs for performing periodic surveillance and maintenance in accordance with plant TS requirements, as well as supporting that specific plant's operating procedures.

2.5 Hardware Quality

2.5.1 Hardware Design

The information should be provided to confirm that the safety-related system equipment is designed to perform the functions for which the equipment is credited in the safety analysis, over the range of environmental conditions postulated for the area in which the equipment is located. Hardware quality and design criteria can be found in the following:

- 10 CFR Part 50, Appendix A, GDC 2, "Design Bases for Protection Against Natural Phenomena," and GDC 4, "Environmental and Dynamic Effects Design Bases"

- 10 CFR Part 50.55a(h), and by reference, either IEEE Std 279, "IEEE Standard: Criteria for Protection Systems for Nuclear Power Generating Stations," or IEEE Std 603-1991
- RG 1.152, "Criteria for Use of Computers in Safety Systems of Nuclear Power Plants," which endorses IEEE Std 7-4.3.2-2003
- RG 1.209, "Guidelines for Environmental Qualification of Safety-Related Computer-Based Instrumentation and Control Systems in Nuclear Power Plants," which endorses IEEE Std 323, "IEEE Standard for Qualifying Class 1E Equipment for Nuclear Power Generating Stations"
- RG 1.180, "Guidelines for Evaluating Electromagnetic and Radio-Frequency Interference in Safety-Related Instrumentation and Control Systems"
- RG 1.75, "Criteria for Independence of Electrical Safety Systems," which endorses IEEE Std 384, "IEEE Standard Criteria for Independence of Class 1E Equipment and Circuits."

The safety-significant CVML system equipment should be designed to perform the functions for which it is credited over the range of environmental conditions for the area in which it is located. It should be further confirmed that the system equipment, including isolation devices and digital equipment, and subject to seismic and environmental qualification requirements, was identified, and design criteria to govern the equipment qualification were established in the application. The hardware design of the CVML system may not involve the manufacturing of components for the specific NPP application, but rather integrate existing (third-party) components for sensing inputs, executing the CVML algorithm, and communicating outputs. Each component of the CVML system must be selected based on specified design requirements, then appropriately tested to confirm that the operational and performance requirements are satisfied—at least at the system level, but also at the component level, if necessary. Design requirements for the CVML system should address the following:

- Input ranges based on the environmental conditions of the target location, with an emphasis on the worst-case conditions for lighting, temperature, humidity, radiation, and pressure; the input sensors should be capable of operating in potentially adverse plant environments, commensurate with the safety function the CVML system was designed to perform.
- Computing resources capable of withstanding the worst-case operating plant and CVML software conditions. This includes considerations of both a localized processing unit and edge computing units located in proximity to the camera. Given the size of typical CVML models, the resources required for the CVML system usually exceed those required by a typical DI&C, and could require more specialized hardware (e.g., GPUs).
- Compatibility of the components within the CVML system should also be a design consideration to minimize downtime due to nuisances (e.g., memory leaks or loss of internal communication buses that require periodic reboot in order to restore operation).
- Outputs capable of providing the necessary information, uncorrupted, within the specified time limits (e.g., specified response time). The output devices and signal characteristics should be unaffected by process or environmental upsets during the time required to perform its function(s).
- Isolation devices (safety-to-non-safety) should be specified as being within the safety boundary and appropriately designed to meet the same operational and environmental requirements as the rest of the safety-related components.
- Maintenance interfaces (e.g., external displays, network devices, and media converters) should be appropriately isolated and non-interfering with the CVML system's safety function. Maintenance tools necessary for long-term maintenance of the CVML system should be available to NPP staff and placed under configuration control to ensure continued and controlled availability over the operational lifetime of the CVML system.

2.5.2 Development Process

As with software QA requirements, a defined hardware development process should be followed that includes management oversight to ensure that quality requirements and project performance measures are tracked and appropriately met. The hardware development process and the quality control methods used during system development should be demonstrated. The information that covers both the development methods used during the design of individual hardware modules and the design of the application-specific system to be used in implementing the safety function should be described. The input information, lifecycle activities, and output information necessary to develop the system should be described. The use of industry standards, including any international standards should be demonstrated. The analysis, review, and test activities to be implemented should be documented and described. The quality control methods used for system development should be consistent with 10 CFR Part 50, Appendix B, and with the criteria of IEEE Std 603-1991.

Information should be provided to confirm that the system equipment and components are designed, developed, fabricated, and tested to quality standards commensurate with the safety significance of the functions to be performed. Design and manufacturing methods and practices should be of sufficient quality to ensure that the CVML system can reliably perform the credited safety functions.

Development of the CVML system should progress according to a defined lifecycle that is part of the overall system development framework. Many different lifecycle models for system, hardware, software, and human factors engineering development are possible, with differences mainly arising in the timing of the various activities and tasks used to produce a high-quality product. Although a particular lifecycle model is not required, the lifecycle activities and tasks should be described, including inputs and outputs that will be implemented in the development of the proposed system. For integrated hardware/software systems, the system development lifecycle may be such that the lifecycle phases (requirements specification, design, implementation) track the final validation and requires integration of the hardware and software into the final system configuration. For the CVML system, the hardware portion may proceed at a faster pace, given that third-party (commercial) components would be utilized. Thus, an important activity would be integration testing to ensure compatibility within the integrated system.

As with software development, the key management activities outlined in the software lifecycle section apply, as well as the requirement for independent design verification: namely, that the design verifier should not have been involved in designing the CVML system.

2.5.3 Equipment Qualification and Appropriate Application of Hardware

Qualification is the process of identifying hazards in the environment in which the I&C systems and equipment may be operating, then conducting tests or analyses (or both) to determine whether the credited safety-significant functions can be reliably performed under the specified service conditions. Therefore, qualification measures should confirm that the CVML system and its equipment can reliably perform the design-basis functions for which they are credited over the range of environmental conditions postulated for the area in which they are located. The following provide regulatory criteria for designing safety-related equipment that can withstand harsh environmental conditions:

- 10 CFR Part 50, Appendix A, GDC 2, “Design Bases for Protection Against Natural Phenomena,” and GDC 4, “Environmental and Dynamic Effects Design Bases”
- 10 CFR Part 50.55a(h), which incorporates, based on the date the construction permit was issued, either IEEE Std 279, “IEEE Standard: Criteria for Protection Systems for Nuclear Power Generating Stations,” Clause 4.4, or IEEE Std 603-1991, Clause 5.4
- RG 1.152, “Criteria for Use of Computers in Safety Systems of Nuclear Power Plants,” which endorses IEEE Std 7-4.3.2-2003, Clause 5.4

- RG 1.209, “Guidelines for Environmental Qualification of Safety-Related Computer-Based Instrumentation and Control Systems in Nuclear Power Plants,” which endorses IEEE Std 323, “IEEE Standard for Qualifying Class 1E Equipment for Nuclear Power Generating Stations,” including five enhancements and exceptions
- RG 1.180, “Guidelines for Evaluating Electromagnetic and Radio-Frequency Interference in Safety-Related Instrumentation and Control Systems.”

The environmental conditions relevant to the application should be defined and described. The normal and extreme ranges of environmental conditions that the system is required to withstand should be specified in accordance with the constraints imposed by the plant design framework. Specified environmental conditions should include the following:

- Temperature, humidity, pressure, and radiation during normal operation and accident conditions
- Conditions imposed by potential hazards external to the system, including seismic conditions, electromagnetic interference, and flooding
- Power supply and heating/ventilation conditions
- Specified environmental qualification of hardware, based on design bases functions; for computer-based systems, this qualification addresses the hardware (including its ability to perform its safety functions under the applicable environmental conditions), the operating system software (if applicable), and representative application software, both integrated in the hardware, based on IEEE Std 7-4.3.2-2003, Clause 5.4, and IEEE Std 603-1991.

For the CVML system, the issues that should be considered when selecting hardware include:

Input requirements

- Signal type
  - Analog, discrete
  - NTSC, PAL, USB, RS232/422/485
  - Ethernet, wireless
  - Custom (proprietary)
- Format compatibility with the CVML algorithm
  - Horizontal resolution
  - Pixel resolution (8-, 10-, 12-bit, etc.)
  - Frame rate
- Environmental compatibility
  - Temperature
  - Humidity
  - Radiation
  - Seismic
  - Sensitivity (lux)

Computing resources

- Structure
  - Type (CPU, field programmable gate array, ASIC)
  - Architecture (ARM, RISC, x86)
  - No. processors (single, multiple)

  - Processor “speed”
    - System clock frequency
    - Floating point operations per second
  - Memory
  - Form factor (laptop, tower, industrial)
- Environmental
  - Temperature
  - Humidity
  - Radiation
  - Seismic (hardened case)

Output requirements

- Signal type (analog, discrete)
- Format (Ethernet, wireless)
- Interfaces (number and types)
  - Monitor
  - Alarm panel
  - USB
  - GPIO
  - Ethernet, wireless
- Environmental
  - Temperature
  - Humidity
  - Radiation

If any of the components require monitoring (e.g., to maintain the validity of the FMEA), the affected parameters should be identified in design documentation so that the operator is provided an appropriate means to monitor the system in the final hardware configuration.

2.5.4 Obsolescence

Obsolescence is used to describe material that remains necessary to operate the plant but is hard to obtain or no longer available. Material obsolescence directly impacts the ability to have the spare parts needed to support plant operations (in any mode of operation). In 2000, the Nuclear Utilities Obsolescence Group was created and developed an obsolescence management program that tracks nuclear parts and the availability thereof. Since then, a number of other vendors have developed programs (e.g., subscription-style software utilities) that help manage the issue to the benefit of the nuclear industry as a whole. Electric Power Research Institute (EPRI) report 1016692 [49] states, “An industry average of approximately 20% of identifiable plant equipment is obsolete (no longer available in the marketplace),” and “Attempts to categorize equipment in a single U.S. plant can result in identification of a population of obsolete equipment nearing 10,000 in number.”

There are no specific safety requirements regarding obsolescence, but the nuclear industry recognizes that it is an issue. As such, procurement specifications often impose requirements that the hardware must be supported throughout the entire operating life of the equipment, including specifying a spare parts list for the CVML system to minimize system interruptions and budget for potential needed hardware (and associated software) upgrades. To ensure timely identification and replacement of components for safety-

related and safety-significant components, complete and accurate bills of material should be provided for the application hardware, as well as sufficient spares for the design life of the system. Within that timeframe, the hardware will necessarily require periodic calibration, functional checks, etc. Additionally, the FMEA may identify credible failures, along with associated calculations for proof-test intervals (e.g., once every 36 months) and component replacement periodicities (e.g., once every 5 years) that necessitate the procurement of spares and the subsequent storage thereof throughout that period.

For commercial-grade hardware—particularly computer systems—provisions should be made to "futureproof" the system against advances in technology that would make it difficult to replace critical hardware components throughout the installed life of the CVML system. One method is to provide long-term support, including periodic upgrades of near-end-of-life components, to ensure continued compatibility of the software with the newer hardware. Ideally, the procurement specification should include requirements for a minimum number of spares (both commissioning and operational) that can be stored on site, as well as requirements for the system supplier/vendor to provide minimum support (over the operating life of the plant), with the minimum number of spares in inventory. The following components may present a risk:

- Cameras
- Computer resources
- Network devices (including any one-way link devices or network diodes).

The discussion of critical characteristics of the system should be provided so that, should components become obsolete, equivalent replacements can be rapidly procured under an effective obsolescence management plan. Any components that are late in product life, rare or customized, or already obsolete should be addressed and should be supplied with the initial delivery in sufficient numbers to ensure adequate operational life while the system is deployed at site.

2.6 Commercial-Grade Dedication

A component (or item) not manufactured under an approved Appendix B program can be used in safety-related applications subject to a 10 CFR Part 50, Appendix B QA program, provided that the component or item meets the 10 CFR Part 21.3 [50] definition of a commercial-grade item (CGI):

- Not subject to design or specification requirements that are unique to nuclear facilities
- Used in applications other than nuclear facilities
- Is to be ordered from the manufacturer/supplier on the basis of specifications as set forth in the manufacturer's published product description (e.g., a catalog).

For safety-related applications, many or all of the components within the CVML system (e.g., a camera) will likely have to be dedicated as basic components in accordance with 10 CFR Part 21. EPRI Topical Report (TR)-106439, "Guideline on Evaluation and Acceptance of Commercial Grade Digital Equipment for Nuclear Safety Applications," as approved by NRC's safety evaluation dated July 17, 1997, describes an acceptable method of performing the necessary evaluation. EPRI TR-106439 provides guidance on evaluating existing commercial computers and software to comply with the criteria of Clause 5.4.2 of IEEE Std 7-4.3.2-2003. The guidance of SRP BTP 7-14 may be applied to evaluations of vendor processes, as described in EPRI TR-106439.

In EPRI NP-5652, "Guideline for the Utilization of Commercial Grade Items in Nuclear Safety-Related Applications" [51], dedication is defined as the action taken by the nuclear industry to utilize a CGI for a safety-related application. This is required by 10 CFR Part 21 [52] (i.e., an item to be dedicated before it can be used as a basic component). EPRI NP-5652 indicates that a CGI dedicated for use in safety-related applications can be shown to be of equivalent quality to a safety-related item purchased as a

basic component. However, to do so, two distinct processes are implemented to provide assurance that the specified item can meet 10 CFR Part 50, Appendix B requirements:

1. A technical evaluation to ensure that requirements for an acceptable item are specified in the procurement document. The technical evaluation process provides a means of specifying the correct requirements for an item in a procurement document.
2. Acceptance methods to reasonably ensure that the item received is the item specified. The acceptance methods for CGI provide reasonable assurance that the item received is the item specified. The technical evaluation, in combination with an appropriate acceptance process, provides assurance that the specified item can meet 10 CFR Part 50, Appendix B requirements.

The dedication process for the digital safety-related system should be described, if applicable. The described dedication process should entail identification of the critical physical, performance, and dependability characteristics necessary to instill adequate confidence that the proposed digital system or component can achieve the safety function. The dedication process should apply to the computer hardware, software, and firmware necessary to accomplish the safety function. For the CVML applications being considered, all third-party components must be dedicated for use, including the input sensors (e.g., cameras), computer resources, and CVML model and algorithm (e.g., if transfer learning is implemented to train the algorithm for the intended task). A benefit of this approach is that, once dedicated, the CVML software can be applied to other uses, with the input data used for retraining having already been controlled, verified, and validated. However, as already discussed, control and continuous revision of the publicly available and used datasets could limit the ability to maintain and update the developed CVML model.

The software dedication process should include an evaluation of both the development process and its implementation to produce the software being dedicated. For commercial-grade software intended for use in safety-related systems, one critical characteristic is the implementation of a high-quality development process. The process of developing the commercial software should be as rigorous as that for software used in safety-related applications; if this cannot be demonstrated, the compensatory measures taken (e.g., extensive operating experience and, if necessary, additional analyses, tests, or inspections) should be developed and described. For CVML software in safety-related applications, dedication of software not produced under an approved 10 CFR Part 50, Appendix B program is needed. Some software QA programs do meet 10 CFR Part 50, Appendix B criteria, but an evaluation of such programs is necessary to identify any gaps. The evaluation would essentially map the developer's program to the requirements (and guidance documents, where appropriate) and identify any gaps.

2.7 System Reliability and Availability

The methods implemented to meet the digital system reliability goals should be developed in accordance with GDC 21, "Protection System Reliability and Testability," of Appendix A, "General Design Criteria for Nuclear Power Plants," to 10 CFR Part 50. However, as explained in RG 1.152, "Criteria for Use of Computers in Safety Systems of Nuclear Power Plants," Revision 3, NRC does not endorse the concept of quantitative reliability goals as the sole means of meeting its regulations for the reliability of digital computers used in safety systems. NRC's acceptance of computer system reliability is based on deterministic criteria for both hardware and software. A quantitative reliability determination using a combination of analysis, testing, and operating experience can add an extra level of confidence in the reliable performance of computer systems.

An important factor in gaining acceptance is to select appropriate reliability metrics and consistently apply them to the CVML system. The metrics can be quantitative, qualitative, or a combination of both. The metrics should be applied correctly and consistently, and be commensurate with the safety significance of the design task of the CVML system.

2.7.1 Hardware Reliability

The methods employed to meet the digital system reliability goals should be developed in accordance with GDC 21 of Appendix A to 10 CFR Part 50. The degree of reliability necessary for the overall DI&C system depends on the safety significance of the system's functions. Therefore, DI&C systems or components should be designed to achieve a reliability level commensurate with the safety significance of the function(s) to be performed. Examples of design attributes for achieving a given level of functional reliability include those related to failure data, fail-safe behavior, independence, redundancy, diversity, failure detection, periodic testing, use of self-diagnostic features, surveillance tests, maintainability, and service life. V&V should be included at appropriate stages of the DI&C system design to confirm that the necessary safety functions were identified and can be reliably performed when called upon. Additional guidance is contained in IEEE Std 603-1991 when applying the associated guidance of IEEE Std 7-4.3.2-2003, Clause 5.15, "Reliability," endorsed by RG 1.152, "Criteria for Use of Computers in Safety Systems of Nuclear Power Plants."

Project or customer requirements for the CVML application may dictate reliability requirements that could be driven by safety requirements. NRC-endorsed standards dictate 99% minimum reliability for safety-related applications involving reactor protection systems. Therefore, for CVML systems intended for such safety-related applications, the initial deployments should involve the development of methods for meeting these high reliability goals. Adequate information for the reviewer should be provided to evaluate whether the proposed CVML system design meets this reliability goal via the use of qualitative or quantitative performance measures or criteria. These performance measures and criteria can be used to optimize goals such as minimizing outage times for repair and reducing the frequency of surveillance. Quantitative reliability is partly determined by the end-to-end function (i.e., from input through the logic processor to the output device). From this perspective, the manufacturer information for each critical hardware component in the loop may suffice for relatively simple safety functions (for series-parallel combinations using given reliability numbers). The initial deployments should consider utilizing methods for calculating probability of failure on demand, mean-time-to-failure, etc., based on the modeling of fault-tolerant systems. Because the use cases of CVML systems in this report (i.e., fire watch and AGRS) would likely not involve protective action (e.g., reactor trip, actuation of emergency core cooling), these metrics may not provide useful information to the reviewer. The selected metrics should be appropriate for the CVML application.

When the use cases are considered in the context of the design principles of redundancy, independence, and separation, qualitative reliability goals may be sufficient to satisfy safe-use criteria. Qualitative analysis, such as for the FMEA, may suffice for demonstrating adequate reliability. Redundant architectures may be necessary to meet availability requirements, as well as the single-failure requirement imposed on the CVML system.

The safety function of the CVML may be composed of multiple SIL-rated hardware components. In these cases, the manufacturers' specified proof-test intervals should be identified (via the FMEA) to ensure that preventive maintenance actions are implemented at the NPP to maintain the validity of the reliability analysis.

2.7.2 Software Reliability

The software design should meet its reliability goals in accordance with GDC 21 of Appendix A to 10 CFR Part 50, and that the qualitative or quantitative performance measures or criteria used to confirm the goals were met for the safety system software. The programmatic controls needed should be identified to address uncertainties and ensure the desired reliability.

There are no accepted methods for quantifying software reliability. There are methods for performing software V&V involving process quality and test metrics that collectively form the basis for accepting software intended for safety-related applications as being sufficiently reliable to perform the design

function(s). Additional guidance is contained in RG 1.152, “Criteria for Use of Computers in Safety Systems of Nuclear Power Plants,” Revision 3, and endorsed standard IEEE Std 7-4.3.2-2003.

Software faults may result from design errors and thus do not reflect the random failure behavior assumed in hardware reliability analyses. Consequently, the reliability analysis may utilize different methods to assess the unreliability introduced by software and hardware. For example, the reliability of CVML systems may be demonstrated on the basis of a combined quantitative-qualitative evaluation, taking into account the complexity of the design, quality of the system, V&V, and testing during the development process over a wide range of input conditions, in addition to feedback from operating experience. Additionally, within the CVML scientific community are various accepted performance measures (Section 2.2) for establishing acceptance criteria for CVML models implemented in software.

For the CVML model, the training, validation, and test databases should include information from the NPP target environment to allow the model to learn the environment without causing overfitting. Techniques are available that prove the CVML models (e.g., DCNNs) can be tricked into false predictions [3,37]. The system’s robustness to adverse conditions should be demonstrated, including low lighting, loud noises, camera vibration, humidity and water spray, and unintended human interaction (e.g., bumping, dropping of tools). For sensitive applications at NPPs, techniques should be applied to ensure that false images (whether corrupted through noisy data or malicious manipulation) cannot cause the model to reflect false negatives, and that the methods applied should be documented to ensure that the model is sufficiently robust.

For those applications in which the input domain can be heavily skewed in one direction, the metrics to establish acceptance thresholds should be carefully selected, and may—specifically in the case of NPP implementations of the CVML model—require multiple metrics. For example, because the consequences of failing to identify a fire (i.e., a false negative) poses a greater risk to the NPP than erroneously declaring a fire (i.e., a false positive), performance metrics centered around the false-negative rate may prove more useful than other metrics. Certain applications may require a balance between metrics for false-negative rates and for overall confidence in the prediction capability of the CVML model. Regardless, the confusion matrix for the CVML model should be provided, and acceptance criteria should be established, with documented justification for numerical thresholds. Objective evidence of satisfying the acceptance criteria would then be generated during V&V testing and retained as quality records. Potential metrics for NPP implementations focused on false-negative rates and overall prediction confidence are as follows (based on the confusion matrix values determined from validation testing):

Threshold metrics

- False negatives
- Accuracy
- Error
- Recall
- F-Measure (harmonic mean)
- Sensitivity (a false-negative indicator)
- Specificity (a true-negative measure)
- G-Mean (geometric mean).

Rank metrics

- Precision-recall curve.

Statistical metrics

- Brier Score.

The threshold metrics result in a number easily comparable to an acceptance threshold, and the rank metrics are plotted against each other to assess the performance of the CVML model. The statistical measure(s) determines the overall confidence in the CVML model predictions. A combination of metrics may be needed to enable the reviewer to make informed decisions as to the acceptability of the CVML system for safety-related or risk-significant applications in an NPP. Ultimately, knowing the end users' requirements and purpose for the CVML system in the NPP will inform the design and testing of the CVML system in order to verify its acceptability. Finally, the selected metrics should be justified in terms of their sufficiency to meet software reliability goals, in combination with the above software QA metrics utilized as part of the software V&V process.

2.7.3 FMEA and Impact on Safety

To demonstrate the integrity of the CVML system, an FMEA of the CVML system should be performed to determine the adverse effects on the system and plant environment. The FMEA scope ranges from input sensors to outputs, including the CVML software algorithm. Credible failures that impact the safety-related function should be resolved through design changes. At minimum, the system and its support systems should provide some indication of failure so that operator actions are not impacted. Potential failure modes may include single random failures, CCFs, etc. Formal analyses of the identified hazards should include a qualitative evaluation whose goals include discovering the fault propagation paths in the systems to determine the root causes of a potential failure mode, and to identify the best ways to minimize the associated risk.

The overall CVML system analysis and the application of the methods used should be documented to demonstrate its reliability are acceptable: (1) the system modeling in the analysis includes the system description, key assumptions, and failure effects, as well as a description of the event sequences following the failure; (2) the overall system design architecture and function allocation support the assumptions in the system modeling. The system modeling should include potential failures of the digital component hardware and software, as well as the design features provided to prevent failures or mitigate or minimize their effects. The acceptability of each failure effect should be justified, through the conducting of a FMEA.

Any potential hazards in the CVML system that could challenge plant safety should be identified, and adequate hazard controls to prevent, eliminate, or mitigate each identified hazard must be provided. The identified hazards, corresponding controls, and techniques used to identify each hazard should be documented. The technique used for hazard identification and control to verify its appropriateness should be described, including any limitations. This should be addressed as part of the overall systematic assessment.

The CVML system should be designed to fail in a safe state, or into a state demonstrated acceptable on some other defined basis. The system—upon detection of inoperable input instruments—should automatically place any safety-related functions associated with the failed instrument(s) into a safe state. The documented FMEA should demonstrate that both hardware failures and software errors detected via self-diagnostics result in safety-related functions being placed into a safe state or leave the protective function in an existing safe state. Computer system hardware failure or software errors should not inhibit manual initiation of protective functions or the performance of preplanned emergency or recovery actions.

This report's use cases for CVML technology in NPPs do not involve any direct protective actions. For these use cases, failure analysis should confirm that the CVML system fails in a known state that does not impair the safety function(s) of connected equipment or delay or confuse operator actions (e.g., frozen reading for AGRS) should the operator resume performance of the design task.

2.8 PRA and Risk Modeling

To modernize NRC regulations, NRC directed its staff to promote, among other approaches, the use of PRA technology in a manner that complements the NRC's deterministic approach and supports the NRC's traditional defense-in-depth philosophy. For example, in Staff Requirements Memorandum to SECY-11-0024, "Use of Risk Insights to Enhance the Safety Focus of Small Modular Reactor Reviews," [53] NRC approved the staff's recommendation to enhance the efficiency and effectiveness of reviewing small modular reactor applications via a design-specific, risk-informed, safety-focused approach.

Currently, the U.S. nuclear industry and NRC staff have not come to an agreement on acceptable methods for modeling DI&C and software in PRAs. Techniques are currently available, incorporation of digital system models into PRA is being performed by utilities and vendors, and new methods continue to be investigated. Given existing techniques and related uncertainties, the question arises as to how an upgrade to a DI&C system involved in CVML-related applications can be modeled so as to illustrate the point at which software CCF becomes risk-significant, along with the degree to which D3 can effectively manage this risk. The answers to these modeling and quantification questions lie in understanding that the I&C for any plant is only one part of a much bigger integrated safety picture, and that the I&C's role in this safety picture is well understood, whether provided by analog, digital systems, or human operator backup. For the current fleet of LWRs, the predominant method of modeling DI&C is to use failure rates mostly derived from hardware failure rates and sensitivity evaluations that bound the impacts of DI&C systems on the overall plant risk. DI&C contributes to plant risk (as measured by CDF) through its potential failures in systems used to mitigate plant events via automated or manual actions, and through its potential spurious failures that initiate events. Both equipment hardware failures and sensitivity studies were evaluated to determine the contribution of I&C failures in mitigating systems to the overall plant risk. The implications of the PRA's sensitivity to the existing I&C are (1) that the PRA results are not sensitive to the modeling and failure probabilities of the mitigating system I&C, but rather the reliability, redundancy, and diversity associated with the mechanical and electrical systems dictate the risk distribution in NPP PRAs; and (2) the bulk of the risk for NPP PRAs results from transient and accident sequences initiated by non-I&C failures and their corresponding frequencies.

From a deterministic licensing perspective, the PRA model on its own fails to address two key issues:

- The well-established methods for modeling DI&C in PRAs do not explicitly model the interactions between the plant system being modeled and the plant physical processes, nor the timing of these interactions
- There are no consensus methods for quantifying the reliability of digital systems due to the difficulty of quantifying software reliability.

The first issue is addressed in plant training simulators, through specific simulated failures that tie I&C failures directly to the plant process behavior and operator interactions to safely manage the plant. The second issue is addressed through extensive simulations to verify the appropriateness of the software under a variety of conditions. Therefore, overall I&C system quantitative reliability goals should support overall plant-level performance objectives, as determined via the PRA results, plant simulations, software verification processes, or other evaluations.

For the use case deployments of the CVML technology in NPPs, there is likely to be minimum plant risk (except perhaps for deployments that impact the FPP). Thus, no additional guidance is given herein beyond that provided in Section 2.7 regarding the reliability of the CVML system and its credible failure modes. This aspect should be coordinated with the NPP organization responsible for performing risk assessments, PRAs, and safety analyses. Through those interactions, the risk that CVML deployment poses to the NPP will be assessed using input from the organization responsible for I&C systems, and the safety analysis report will be updated accordingly.

2.9 Other Topics

2.9.1 Human Factors

Some DI&C equipment modifications may involve human factors engineering (HFE) considerations (e.g., HFE analyses and design processes). Previous sections discussed the requirements related to the various means by which operators interact with the CVML system, such as system maintenance interfaces and the outputs to operator displays used in the performance of manual actions. If the design affects indications used by the operator for manual control, the modification effect on the operator's ability to implement manual actions should be developed, as per IEEE Std 603-199. The interface and controls associated with status and bypass indications should be identified, per IEEE Std 603-1991. If the design affects indications used by the operator for manual control or status/bypass indications, the modifications should support the operator's ability to use the indications, per IEEE Std 603-1991. An HFE safety evaluation may need to be performed in accordance with SRP, Chapter 18, "Human Factors Engineering;" NUREG-0711, "Human Factors Engineering Program Review Model;" and NUREG-1764, "Guidance for the Review of Changes to Human Actions," in close coordination with the I&C evaluation described in Chapter 7.

Human acceptance of the CVML system should be considered. The trustworthiness of CVML technology must be considered from at least two perspectives [54] First, from the operator perspective, because the CVML algorithm is treated as a black box, the quality of the algorithm's decisions is difficult to determine until plant conditions are perhaps beyond recovery; thus, abnormal operating conditions may ensue, placing undue stress on the operators to recover the plant. One means of improving trustworthiness is explainability, demonstrated in [41], since information from the NN's intermediate layers can be extracted for meaningful presentation to the operators. Deployments of CVML systems could present information based on the intermediate layers, enabling the operators to determine the quality of the CVML decision making [54]. Second, from the regulator perspective, having so few examples of CVML systems currently deployed in NPPs generates very little operating experience. Thus, regulatory oversight of initial deployments of this technology may be onerous, requiring that high expectations for the information provided be met prior to granting approval.

2.9.2 Operations/Operating Environment

A system is considered operable once it can perform its intended function (e.g., detect fire in the target location). This definition of operability includes the principle that a system can perform its specified safety functions only when all its necessary support systems can perform their related support functions. As part of obtaining an operating license, utility owners/operators are required by 10 CFR Part 50.34 [55] to establish programs that ensure safe operations through managerial and administrative controls. The plant's FSAR lays out the controls in accordance with the utility's approved 10 CFR Part 50, Appendix B program. The plans should address the conduct of normal operations, including maintenance, surveillance, and periodic testing of the plant's infrastructure and critical equipment (i.e., systems and components). The plans should also address coping with emergencies and preparing TS in accordance with 10 CFR Part 50.36 [56].

The operability assessments should be conducted and prompt corrective action in inoperable, degraded, or nonconforming conditions should be identified. Any system dependent on an inoperable support system (e.g., power, communication) would also be declared inoperable. The process of ensuring operability is continuous and consists of verifying operability via surveillance activities and formal determinations of operability whenever a verification or other indication calls into question the ability of a structure, system, or component to perform its specified function. Prompt action should be taken by the licensee any time a safety-important structure, system, or component is deemed inoperable.

One important aspect of the regulation is the requirement to develop the TS for establishing and maintaining critical plant parameters (e.g., safety limits for trip setpoints) and establishing compensatory

measures when safety-related systems or components are removed from service, whether for planned maintenance or due to failure. Changes to systems or components may impact the TS, and thus may be accompanied by a change to the plant TS. The plant TS controls critical parameters for safety-related DI&C systems, as well as certain compensatory measures for specific plant configurations.

During the operating life of the NPP, the utility owners/operators are further required by 10 CFR Part 50.65 [57] to monitor the effectiveness of the plant maintenance against established performance goals (e.g., preventive and corrective maintenance). This includes monitoring the condition of systems to ensure their operability and availability to perform their design functions. This maintenance rule extends to non-safety systems when those systems are relied on to mitigate accidents or are used as part of emergency operations, when failure of the non-safety-related system could prevent safety-related systems from performing their safety functions, or when failure of the non-safety-related system could cause an emergency plant shutdown.

The impact on the plant's operational mode (i.e., configuration), as well as any new or modified maintenance activities that the change will impose should be addressed. As a result of the safety requirement for monitoring maintenance, the plant activities should be preplanned and strictly controlled, particularly during plant outages. This would require many hours of preparation and copious amounts of documentation to serve as objective evidence that the maintenance was performed adequately and the equipment was left in an operable state. Given the complexity of NPP environments, plant configurations before, during, and after maintenance must be considered when installing a new or replacement system, potentially resulting in temporary configurations that require compensatory measures (e.g., under the plant's FPP). The temporary plant configurations that, during maintenance, could impact the new or modified system should be developed, including interfacing systems (e.g., communications network), environmental changes (e.g., lighting, barriers, human traffic), and operational changes (e.g., maintenance bypass, removal from service).

During normal operations, the NPP generally features a stable environment in which the visual view of equipment is predominantly fixed under constant (unchanging) optical conditions. However, variations in plant configurations from one operating mode to another can be dramatic, particularly during shutdown or maintenance outages in which temporary equipment might be staged and above-normal human traffic present. This includes conditions in which the physical environment has been changed due to staged repair equipment, scaffolding, and worker high-traffic conditions. The CVML system should be capable of performing its task in such temporary conditions—particularly, emergency situations or abnormal operations in which relatively short-lived but harsh conditions exist. Ensuring that the CVML system is robust to relatively harsh and dynamic operating environments—even to perform only relatively simple tasks—will provide significant operating experience and gain the trust of operators and regulators alike. Options for addressing those condition modes include training the CVML model to recognize the different operating modes, having a dedicated CVML model for each operating mode, or teaching the NPP maintenance staff how to train the CVML models to accommodate new conditions.

Given the rapid, recent advancement of CVML, there may be a lack of qualified staff with the expertise to perform the various activities required to sustain the system operability. NPP staff from design/engineering, operations, and maintenance organizations should be trained on the CVML technology (and specific configurations deployed at the NPP) to prevent over-reliance on specialized third-party expertise to deploy, operate, diagnose, and repair the CVML system. This should include providing all the necessary tools required for configuring and modifying/repairing the CVML software algorithm (e.g., via retraining the system). This necessarily requires user manuals for installation, operation, and maintenance of the complete CVML system.

2.9.3 Cybersecurity

As digital technology continues to progress at an ever-faster pace, network architectures and data storage methods create new challenges with regard to cybersecurity for critical infrastructure. A DI&C

system should meet the requirements of 10 CFR Part 73.54 [58], "Protection of Digital Computer and Communication Systems and Networks," which addresses cybersecurity for digital assets, including those systems used to perform functions important to safety, security, and emergency preparedness. Secure development and operating environment vulnerability assessments should be performed in accordance with the guidance in RG 1.152.

The CVML system should adhere to the cybersecurity protection program as discussed in NEI 08-09, "Cyber Security Plan for Nuclear Power Reactors," [59] and NRC RG 5.71, "Cyber Security Programs for Nuclear Facilities" [60], to ensure adequate protection of critical digital assets. The proposed CVML architecture should be reviewed against the criteria of RG 5.71 and NEI 08-09 for potential vulnerabilities.

There are multiple cybersecurity considerations when using the CVML system in the nuclear industry. First, there is publicly available source code. As already discussed, many advanced CVML models are based on publicly available models that utilize massive databases compiled over many years in the public domain. This may present significant challenges to adapting the technology to the secure operating environment of an NPP. Additionally, the continuous learning of the CVML models (e.g., the weighting factors in the convolution layer of the DCNN) enables the software to perform autonomous software changes that could be construed as software configuration changes that fall outside the expected software V&V process, as defined in the NRC and industry software design guidance. Thus, it could be susceptible to cybersecurity attacks. If the model must continue to learn following the installation of the system (as more images or videos are added to the database from multiple plants), a connection between the system and the main network system, which could be located offsite (e.g., on a cloud) from the plant, would be needed for continuous model tuning or to access an image database. Rigorous analysis of the security of the interface to the image database identified mitigating measures on the NPP side of the network to ensure that data could not be manipulated to teach the system to incorrectly read the gauge images, for example. The analysis should leverage the hardware and software FMEAs discussed previously, as well as the cybersecurity program requirements to identify and implement protective measures.

3. SUMMARY OF CVML-SPECIFIC CONSIDERATIONS

From this study, the following major CVML-specific characteristics atypical of traditional DI&C must be given dedicated consideration via suggested solutions (tailored for each requirement). A summary is presented in Table 1 and discussed herein. This summary does not represent a comprehensive list of all the characteristics and the associated considerations as they often correlate to each other. Instead, it aims to provide a context of the type of considerations needed with such a technology.

1. CVML models often utilize open-source datasets and feature extraction engines or models. The impact of this can be summarized as:
   - It is not always possible to determine the level of overlap among open-source datasets. Open-source models could use similar fundamental concepts. This could impact the independence of the developed CVML models when used to demonstrate D3. This implies that the data or model could make the CVML system susceptible to CCF. Also, the independence (or lack thereof) of the data or model will have to be taken into consideration during the design verification process, since the overlap could overestimate the software V&V performance results. To address this consideration:
     - Methods may be needed to ensure adequate diversity of the data (e.g., using different sets) when testing models against defined acceptance criteria.
     - Tools may be needed to identify overlapping images/video in datasets.
     - Methods to create independent datasets may be needed, such as using a generative adversarial network to generate datasets that are guaranteed unique, augmentation (e.g., by rotating an object in an image frame) to increase the size of the training dataset, or development of isolated benchmarking datasets that are independently and manually assembled.
     - The underlying concept used by the CVML models need to be evaluated for independence. Various forms of CVML ANNs could be used to ensure model independence.
     - Validation should rely on using multiple different packages or tools (e.g., using PyTorch versus using TensorFlow), in addition to using different environments.
   - Frequent updates to the open-source datasets and models are important for advancing the performance of the CVML model, but are not controlled by the CVML developer, thus affecting the configuration control over design inputs as well as the software CM processes. The lack of control and the open-source nature of the dataset and models also introduces a cybersecurity concern. The CGD process for those models might also be infeasible via conventional means used for DI&C. If periodic updates are necessary, it could introduce a maintainability burden beyond that usually expected for DI&C. To address this consideration:
     - The original images/videos (used for training, validation, and testing) and models (used for feature extraction) should be placed under some form of configuration control (e.g., create a controlled local copy), whereby the initial version is recorded and protected from inadvertent, incorrect, and unauthorized changes. This ensures traceability to a known source, as well as tracking of all changes.
     - New input training data should be thoroughly vetted and documented for appropriateness to the application and the diversity and validation requirements. The updated training dataset should be carefully examined and filtered via the same metrics used in the initial design to ensure it matches the evaluation process and quality of the initial design input.
     - Methods may be needed to identify added images/videos so that any revision of the input data files is recognized before being used for model training.

Table 1. Summary of the main characteristics of CVML and potential considerations (in dark gray) when used in a safety-related application.

Characteristic/Consideration	Independence	Defense in Depth	CCF	V&V	QA	Configuration Control	Cyber Security	CGD	Maintainability	Traceability	Design Control	Repeatability	Deterministic Nature	Explinability	Reliability	FMEA	Simplicity	Justification	Trustworthiness
Open-source data and model																			
Frequent updates to source																			
Massive amounts of data																			
Periodic training																			
Probabilistic and stochastic																			
Various performance metrics																			
Incomprehensible to reviewers																			
Inherited bias																			
Non-systematic approach																			
Robustness to new conditions																			
Special skillset																			

- CGD needs to be performed for commonly used models. The CGD evaluation would essentially map the model developer's program to the safety requirements (and guidance documents, where appropriate) and identify any gaps.

2. A CVML model requires massive amounts of data for training and validating. The impact of this can be summarized as:

   - Manual QA of the design input dataset is time consuming and often impractical without some form of automated means. Traceability of data issues and failure causes is also challenging. To address this consideration:
     - Tools may be needed to identify added images/videos and ensure that any revision of the input data files are flagged for QA review.
   - It could be challenging to thoroughly vet and document the revisions (for appropriateness to the application) of the design, thus impacting the design control and software CM. To address this consideration:
     - The data may need to be examined via automated metrics (e.g., for measuring image quality and feature visibility) to ensure that it is of satisfactory quality.
     - Tools may be needed for automated dataset verification.
   - More capable and complex hardware design is needed to handle the large amount of data and preprocessing needs. Dedicated hardware (e.g., ML-dedicated GPUs) for CVML model analysis, that is not typically used for DI&C, would need to be qualified. To address this consideration:
     - CGD methods are needed for commonly used hardware in CVML applications.

3. Periodic training of the CVML models may be necessary to enhance the CVML system performance, especially as it is deployed and limitations are identified. The impact of this can be summarized as:

   - The CVML model could periodically adjust its internal parameters to improve its output performance. This capability may be required to maintain the model and adapt it to operational changes and findings but may not conform to the current guidance related to design control and CM processes. It also introduces additional sources of risk that must be evaluated. To prevent auto-configuration, a manual periodic training process can be used, but this is time consuming, and can impact the maintainability of the CVML technology. Additionally, autonomous software changes could have cyber vulnerabilities (especially if connected to a cloud that hosts the dataset or models). To address this consideration:
     - If a CVML model can be corrupted by data when in learning mode, that mode should be disabled (i.e., batch learning only).
     - To reduce overall risk, a diverse function could take the form of a completely different system, such as a thermal camera for a fire watch CVML system.
     - Rigorous analysis of the security of the image database and of the interface to the dataset on the NPP side of the network may need to be performed to ensure that data are protected from unauthorized access and manipulation. The analysis should leverage the hardware and software FMEAs discussed previously, as well as the cybersecurity program requirements to identify and implement protective measures.

4. CVML models use probabilistic thresholding functions and are dependent on training-specific stochastics parameters. The model could be sensitive to minor design decisions or changes (especially in cases of overfitting) and to small variations in the operational input, thus generating unpredictable output. The impact of this can be summarized as:

   - This introduces unrepeatability (i.e., unreproducible and unpredictable performance concerns) and trust issues, and contradicts the deterministic behavior requirement. To address this consideration:
     - Methods to test and protect against overfitting (e.g., regularization) should be used to ensure the model is not learning the structure of the training data, but rather the modeled process.
     - Sensitivity analysis should be performed to demonstrate the CVML's ability to generate consistent results after retraining with an identical set of images/videos or following a design change.
     - The results of acceptance testing should allow the assessment of the CVML probabilistic aspect and perhaps establish an acceptance threshold.
   - The design verification process does not always yield the same output for the same input, an important requirement for the design verification and software V&V processes. To address this consideration:
     - The V&V process should leverage predesigned test cases with well-defined acceptance criteria for documenting the CVML system's performance.
     - Explainability methods can be used to reveal the CVML model's behavior in decision making.
   - Given the software size of the CVML model, corruption of the CVML model or image, whether by corruption of the executing program or a hardware glitch that causes the CVML model to initiate a processing error is more likely than in typical DI&C and can significantly impact its performance—a key aspect to consider in FMEA and any PRA-related analysis. To address this consideration:
     - Fault detection and diagnosis must emphasize the likelihood of a CVML data corruption (as a mode of failure) and account for this possibility in the software reliability evaluation.
   - CVML model performance represents a compromise between true positives/negatives and false positives/negatives. The performance metrics used can be misleading. To address this consideration:
     - The fault-tolerant performance must be carefully analyzed using FMEA, along with the types of fault detection and diagnosis needed for each application, supplemented by the acceptance criteria used in software V&V and quantitative/qualitative metrics applied in the software reliability evaluations.

5. If the CVML system is based on DCNNs, DCNNs are particularly efficient as image classifiers but often consist of a mathematical model that is incomprehensible to reviewers, and thus are considered a form of black box. The impact of this can be summarized as:

   - It contradicts the design explainability and simplicity, and impacts the depth of the design verification process, since the verification and review may require analyzing the CVML model and getting inside the black box—something not always feasible. To address this consideration:
     - A safety benefit should be used to justify any added complexity, including the use of DCNNs.
     - Given the unique, unconventional nature of CVML systems (as compared to typical DI&C), a significant amount of technical information should be provided for a first-of-a-kind implementation at an NPP.

- Inherited bias in the model may not be detected via model examination or through V&V, thus impacting the independent design verification and software V&V processes and possibly cause a CCF. To address this consideration:
  - A detailed description of the internal CVML model functions should be provided (to the extent possible) to enable a certain level of CVML model evaluation (i.e., getting inside the black box) and give a better understanding of, among other things, numerical errors (and their contribution to bias) and potential failure modes.
  - The inability to review the CVML model should be compensated for by using explainable and transparent models. Some implementations extract data from intermediate layers to achieve those objectives.
- Major DCNN design decisions (e.g., the model architecture used, the size of the DCNN, and the type of functions used within the layer) rely on non-systematic criteria and architectures that are either based on the literature or on experimental results. However, the design verification requires justification for every design decision, with sufficient evidence to support each decision. Also, an independent verification might not use the same DCNN architecture that was developed in the design process. To address this consideration:
  - Systematic design decisions involving tradeoffs should drive the design process, get documented, and be provided when available.
  - The design and verification process should use different CVML models with different architectures (and datasets), the outputs of which could be voted on and compared prior to conclusion or action.
- The trustworthiness of CVML technology is of concern from the operator and evaluator perspectives. To address this consideration:
  - Dedicated means of enhancing the CVML technology's trustworthiness to operators and evaluators must be implemented to compensate for the black box nature of—and lack of operational experience for—the CVML model. A means to reveal aspects of the decision making must be created to improve the perception of the CVML technology from a human factors standpoint and foster greater usage of it.
  - It is necessary to emphasize that the CVML system for any plant is only one part of a much bigger integrated safety picture, and that its role in this safety picture is well understood.

6. Ensuring that the CVML system is robust in relatively harsh and dynamic operating environments depends on the data used for training. The impact of this can be summarized as:
   - The robustness of CVML models for rare events is of concern, especially for unexpected conditions that are relevant to the application but not reflected in the training process. The results can be unpredictable, resulting in CCF and trust issues, which impact the deterministic behavior requirement. This also impacts software reliability and should be accounted for in the FMEA, as well as addressed when considering the operating environment. To address this consideration:
     - The training of the CVML model should consider the different operating modes and adverse conditions including low lighting, loud noises, camera vibration, humidity and water spray, and unintended human interaction (e.g., bumping, dropping of tools).
     - Dedicated techniques should be used to assess the robustness of the CVML model (e.g., attempt to "trick" the CVML model into false predictions).
     - A dedicated CVML model should be developed for each operating mode.
     - The NPP maintenance staff should be able to train the CVML models when needed to accommodate new conditions.
     - Varied performance measures need to be evaluated to determine the suitable combinations of measures to ensure acceptability of the CVML mode, especially in detecting rare events.

7. Given the rapid and recent advancement of CVML technology, there may be a lack of qualified staff with the expertise to perform the various activities required by the safety framework. The impact of this can be summarized as:
    - The NPP could be over-reliant on third parties when ensuring that the CVML system is compliant with the applicable safety requirements and guidance. While this impacts most of the requirements discussed in this report, it has a special impact on maintainability, as well as the capability of the plant staff to perform fault detection/diagnostics and understand the cause, perform CM-related changes, and operate the system. To address this consideration:
        - NPP staff from design/engineering, operations, and maintenance organizations should be trained on the CVML technology (and specific configurations deployed at the NPP).

4. REFERENCES

1. Lybeck, N. J., Thomas, K. D., and Primer, C. A. 2018. "Plant Modernization Technical Program Plan." INL/EXT-13-28055 Revision 8, Idaho National Laboratory.
2. Hallbert, B., Leonard, K.,Farmer, M., Primer, C. A.,Szilard, R. 2018. "Light Water Reactor Sustainability Program Integrated Program Plan," INL/EXT-11-23452 Revision 6, Idaho National Laboratory.
3. Suman, S. 2021. "Artificial intelligence in nuclear industry: Chimera or solution?" *Journal of Cleaner Production*, 278(1):124022. doi: 10.1016/j.jclepro.2020.124022.
4. U.S. NRC, "10 CFR PART 100—Reactor Site Criteria," Accessed September 2021: https://www.nrc.gov/reading-rm/doc-collections/cfr/part100/full-text.html.
5. U.S. NRC, "Appendix B to Part 50—Quality Assurance Criteria for Nuclear Power Plants and Fuel Reprocessing Plants," Accessed September 2021: https://www.nrc.gov/reading-rm/doc-collections/cfr/part050/part050-appb.html.
6. U.S. NRC. 2018. "Digital Instrumentation and Controls DI&C-ISG-06 Interim Staff Guidance," Revision 2. https://www.nrc.gov/docs/ML1826/ML18269A259.pdf.
7. U.S. NRC, "10 CFR 50.48—Fire protection," Accessed September 2021: https://www.nrc.gov/reading-rm/doc-collections/cfr/part050/part050-0048.html.
8. U.S. NRC, "10 CFR 50—Appendix A to Part 50—General Design Criteria for Nuclear Power Plants," Accessed September 2021: https://www.nrc.gov/reading-rm/doc-collections/cfr/part050/part050-appa.html.
9. U.S. NRC, "Part 50—Domestic Licensing of Production and Utilization Facilities," Accessed September 2021: https://www.nrc.gov/reading-rm/doc-collections/cfr/part050/index.html.
10. U.S. NRC, "10 CFR 50—Appendix R to Part 50—Fire Protection Program for Nuclear Power Facilities Operating Prior to January 1," Accessed September 2021: https://www.nrc.gov/reading-rm/doc-collections/cfr/part050/part050-appr.html.
11. U.S. NRC, "10 CFR 50.12—Specific exemptions," Accessed September 2021: https://www.nrc.gov/reading-rm/doc-collections/cfr/part050/part050-0012.html.
12. U.S. NRC, "10 CFR 50.59—Changes, tests and experiments," Accessed September 2021: https://www.nrc.gov/reading-rm/doc-collections/cfr/part050/part050-0059.html.
13. U.S. NRC, "10 CFR 50.71—Maintenance of records, making of reports," Accessed September 2021: https://www.nrc.gov/reading-rm/doc-collections/cfr/part050/part050-0071.html.
14. U.S. NRC, "Regulatory Guide 1.189—Fire Protection for Nuclear Power Plants" Revision 2, October 2009. https://www.nrc.gov/docs/ML0925/ML092580550.pdf.
15. NFPA. 2020. "NFPA 805—Performance-Based Standard for Fire Protection for Light Water Reactor Electric Generating Plants," https://www.nfpa.org/codes-and-standards/all-codes-and-standards/list-of-codes-and-standards/detail?code=805.
16. Nuclear Energy Institute, "NEI 04-02—Guidance for Implementing A Risk-Informed, Performance-Based Fire Protection Program Under 10 CFR 50.48(c)," September 2005. https://www.nrc.gov/docs/ML0525/ML052590476.pdf.
17. U.S. NRC, "Regulatory Guide 1.205—Risk-Informed, Performance-Based Fire Protection for Existing Light-Water Nuclear Power Plants," May 2006. https://www.nrc.gov/docs/ML0611/ML061100174.pdf.

18. U.S. NRC, "10 CFR 50.55a—Codes and standards," Accessed September 2021: https://www.nrc.gov/reading-rm/doc-collections/cfr/part050/part050-0055a.html.

19. IEEE, 1991. "IEEE 603-1991—IEEE Standard Criteria for Safety Systems for Nuclear Power Generating Stations," https://standards.ieee.org/standard/603-1991.html.

20. IEEE, 2003. "7-4.3.2-2003–IEEE Standard Criteria for Digital Computers in Safety Systems of Nuclear Power Generating Stations." https://ieeexplore.ieee.org/document/1263353.

21. IEEE, 1998. "IEEE 603-1998–IEEE Standard Criteria for Safety Systems for Nuclear Power Generating Stations," https://standards.ieee.org/standard/603-1998.html.

22. IEEE, 2004. "IEEE 1012-2004—IEEE Standard for Software Verification and Validation." https://standards.ieee.org/standard/1012-2004.html.

23. U.S. NRC, 2013. "Regulatory Guide 1.168—Verification, Validation, Reviews, And Audits for Digital Computer Software Used in Safety Systems of Nuclear Power Plants," Revision 2. https://www.nrc.gov/docs/ML1307/ML13073A210.pdf.

24. U.S. NRC, 2011. "Regulatory Guide 1.152—Criteria for Use of Computers in Safety Systems of Nuclear Power Plants," Revision 3. https://www.nrc.gov/docs/ML1028/ML102870022.pdf.

25. U.S. NRC, "Standard Review Plan for the Review of Safety Analysis Reports for Nuclear Power Plants: LWR Edition," (NUREG-0800, Formerly issued as NUREG-75/087), Accessed September 2021: https://www.nrc.gov/reading-rm/doc-collections/nuregs/staff/sr0800/index.html.

26. U.S. NRC, 2016. "Regulatory Standard Review Plan- Appendix 7.0-A Review Process for Digital Instrumentation and Control Systems," Revision 6. https://www.nrc.gov/docs/ML1601/ML16019A085.pdf.

27. U.S. NRC, 2007. "Branch Technical Position 7-14 - Guidance on Software Reviews for Digital Computer-Based Instrumentation and Control Systems," Revision 5. https://www.nrc.gov/docs/ML0706/ML070670183.pdf.

28. Wikipedia, "Outline of Machine Learning," Accessed September 2021: https://en.wikipedia.org/wiki/Outline_of_machine_learning#Machine_learning_algorithms.

29. Elias, R., 2017. "Artificial Intelligence terminologies," Accessed September 2021: https://medium.com/machine-learning-world/artificial-intelligence-terminologies-260f1d6d299f.

30. SharperAI, Blog. "A Brief Taxonomy of AI," Accessed September 2021: https://www.sharper.ai/taxonomy-ai/.

31. Rawat, W. and Wang, Z. 2017. "Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review," *Neural Computation*, 29(9):2352-2449. doi: 10.1162/neco_a_00990.

32. Al Rashdan, A., Griffel, M., and Powell, L. 2019. "Automating Fire watch in Industrial Environments through Machine Learning-Enabled Visual Monitoring," INL/EXT-19-55703, Idaho National Laboratory.

33. Hochreiter, S. and Schmidhuber, J. 1997. "Long Short-Term Memory," *Neural Computation*, 9(8):1735-1780. doi: 10.1162/neco.1997.9.8.1735.

34. Abu-El-Haija, S., Kothari, N., Lee, J., Natsev, P., Toderici, G., Varadarajan, B. and Vijayanarasimhan, S., 2016. Youtube-8m: A large-scale video classification benchmark. arXiv preprint arXiv:1609.08675.

35. Huang, J., et al. 2017. "Speed/accuracy trade-offs for modern convolutional object detectors," arXiv: 1611.10012. https://arxiv.org/abs/1611.10012.

36. Dunnings, A. and Breckon, T. P. 2018. “Experimentally Defined Convolutional Neural Network Architecture Variants for Non-Temporal Real-Time Fire Detection,” 25th IEEE International Conference on Image Processing.

37. Muhammad, K., et al. 2018. “Convolutional Neural Networks based Fire Detection in Surveillance Videos,” IEEE Access.

38. Redmon, J. and Farhadi, A., 2018. “YOLOv3: An Incremental Improvement,” University of Washington, arXiv: 1804.02767.

39. Wong, A. and Kamel, M. 2009. “Classification of imbalanced data: a review,” *International Journal of Pattern Recognition and Artificial Intelligence*, 23(4):687-719. doi: 10.1142/S0218001409007326.

40. Muhammed, K., et al. 2018. “Efficient Fire Detection for Uncertain Surveillance Environment,” *IEEE Transactions on Industrial Informatics*, 15(5).

41. Muhammed, K., et al. 2019. “Efficient Deep CNN-Based Fire Detection and Localization in Video Surveillance Applications,” *IEEE Transactions on Industrial Informatics*, 49(7):1419-1434. doi: 10.1109/TSMC.2018.2830099.

42. U.S. NRC, “Regulatory Guide 1.53 Application of The Single-Failure Criterion to Safety Systems,” Revision 2, November, 2003: https://www.nrc.gov/docs/ML0332/ML033220006.pdf.

43. IEEE, “IEEE 379-2000—IEIEEE Standard Application of the Single-Failure Criterion to Nuclear Power Generating Station Safety Systems 2000,” https://ieeexplore.ieee.org/document/914366.

44. U.S. NRC, “Regulatory Guide 1.173 Developing Software Life-Cycle Processes for Digital Computer Software Used in Safety Systems of Nuclear Power PLANTS,” Revision 1, July 2013: https://www.nrc.gov/docs/ML1300/ML13009A190.pdf.

45. ASME, “Quality Assurance Requirements for Nuclear Facility Applications 2020,” https://www.asme.org/codes-standards/find-codes-standards/nqa-1-quality-assurance-requirements-nuclear-facility-applications.

46. U.S. NRC, “Regulatory Guide 1.28 Quality Assurance Program Criteria (Design and Construction),” Revision 5, October 2017: https://www.nrc.gov/docs/ML1720/ML17207A293.pdf.

47. U.S. NRC, “NUREG/CR-6101 Software Reliability and Safety in Nuclear Reactor Protection Systems,” June 1993: https://www.nrc.gov/reading-rm/doc-collections/nuregs/contract/cr6101/cr6101.pdf.

48. U.S. NRC, “Regulatory Guide 1.169 Configuration Management Plans for Digital Computer Software Used in Safety Systems of Nuclear Power Plants,” July 2013: https://www.nrc.gov/docs/ML1235/ML12355A642.pdf.

49. EPRI. 2008. “Plant Support Engineering: Obsolescence Management,” 1016692: https://www.epri.com/research/products/1016692.

50. U.S. NRC, “10 CFR 21.3—Definitions,” Accessed September 2021: https://www.nrc.gov/reading-rm/doc-collections/cfr/part021/part021-0003.html.

51. EPRI, “Plant Engineering: Guideline for the Acceptance of Commercial-Grade Items in Nuclear Safety-Related Applications,” Revision 1 to EPRI NP-5652 and TR-102260, September, 2014: https://www.epri.com/research/products/000000003002002982.

52. U.S. NRC, “10 CFR 21—Reporting of Defects and Noncompliance,” Accessed September, 2021: https://www.nrc.gov/reading-rm/doc-collections/cfr/part021/index.html.

53. U.S. NRC, SECY-11-0024, “Use Of Risk Insights To Enhance The Safety Focus Of Small Modular Reactor Reviews”, Last accessed September, 2021, link: https://www.nrc.gov/docs/ML1107/ML110730424.pdf.
54. Coble, J., Idaho National Laboratory Symposium on Machine Learning-Artificial Intelligence 4.0: Trustworthy AI, “Explainable AI to Support Operations and Maintenance at Nuclear Power Plants,” University of Tennessee-Knoxville, February 2021.
55. U.S. NRC, “10 CFR 50.34—Contents of applications; technical information,” Accessed September, 2021: https://www.nrc.gov/reading-rm/doc-collections/cfr/part050/part050-0034.html.
56. U.S. NRC, “10 CFR 50.36—Technical specifications,” Accessed September 2021: https://www.nrc.gov/reading-rm/doc-collections/cfr/part050/part050-0036.html.
57. U.S. NRC, “10 CFR 50.65—Requirements for monitoring the effectiveness of maintenance at nuclear power plants,” Accessed September 2021: https://www.nrc.gov/reading-rm/doc-collections/cfr/part050/part050-0065.html.
58. U.S. NRC, “10 CFR 73.54—Protection of digital computer and communication systems and networks,” Accessed September 2021: https://www.nrc.gov/reading-rm/doc-collections/cfr/part073/part073-0054.html.
59. NEI, “NEI 08-09 Cyber Security Plan for Nuclear Power Reactors,” Revision 6, April 2010: https://www.nrc.gov/docs/ML1011/ML101180437.pdf.
60. U.S. NRC, “Regulatory Guide 5.71 Cyber Security Programs for Nuclear Facilities,” January 2010: https://www.nrc.gov/docs/ml0903/ml090340159.pdf.