JOURNAL OF INTERNATIONAL CRISIS AND RISK COMMUNICATION RESEARCH
ISSN: 2576-0017
2021, VOL 4, NO S4

Enhancing Chip Performance Through Predictive Analytics and Automated Design Verification

Botlagunta Preethish Nandan

Senior Lead Business Intelligence, preethishnananbotlagunta@gmail.com , ORCID ID : 0009-0008-3617-8149

Abstract

Advancements in chip architecture and process technology have resulted in continued shrinkage of chip areas while integrating more and more functionality. Moreover, these chips are increasingly operating at higher performance and speed levels, often approaching thermal, power and reliability limits. This technical trend calls for new methodologies in the design verification of chips, especially in Exhaustive Detection of non-robust chips, since the costs of not detecting these chips early in their design life are becoming prohibitively expensive. This paper describes a systematic design methodology and new software tools that enable efficient use of predictive modelling and predictive analytics within the design cycle during implementation, compilation and verification. The predictive modelling relies on a combination of procedural and neural network based analytical models to estimate performance metrics early in the design cycle, preferably during synthesis. Then these predictive models are dynamically modified and further optimized during design compilation and data placement verification using a combination of predictive analytics and automated design convergence techniques to ensure chip quality-of-results.

In order to address the issue of Exhaustive Detection of thermal, power and thus reliability non-robust chips, we propose the design methodology and tools for verification that are crucial to the success of new Robust Design concepts, such as Adaptive Speed and Adaptive Voltage designs which are known to alleviate thermal and power non-robustness during normal chip operation. This paper illustrates the methodology with both analog and digital examples and discusses its implementation in a number of tools and systems.

Key Words : Enhancing chip performance, predictive analytics, automated design verification, semiconductor optimization, machine learning, performance tuning, hardware acceleration, real-time monitoring, error detection, design automation, verification tools, fault prediction, reliability analysis, design efficiency, AI-driven testing, chip architecture, power optimization, latency reduction, data-driven modeling, yield improvement, system-on-chip (SoC), computational efficiency, fabrication accuracy, intelligent diagnostics, performance metrics, predictive modeling.

1. Introduction

With the increasing use of telemetry in modern System-on-Chip designs, many chips have ramped into production with custom logging/monitoring components. Test points have recently been added during the test phase in existing designs to track potential chip performance degeneration modes caused by various usage and environmental conditions. These test points are becoming more powerful diagnostic tools every day with the help of increasingly sophisticated onboard software and high-speed internal buses, enabling many chips to upload log data in real-time to the cloud whenever there is a performance anomaly. This availability of such large and abnormal-

event-focused log data provides us unique opportunities to enhance chip performance and build better chips. We provide an overview of how predictive analytics based on chip telemetry data can be harnessed to enhance SoC performance. These analytics models are used to perform real-time as well as post-silicon predictive analytics and are also used for closed-loop chip design verification during production test phases, thereby improving overall SoC performance.

In the first part of this essay, we review the ability of chip telemetry data to predict in-chip performance issues such as high CPU frequency, low CPU frequency, high GPU frequency, high memory controller queue, etc. Additionally, we also explore the potential of chip telemetry data to predict the following chip performance-related metrics: dram_data_pass_fail_frac, dram_timing_pass_fail_frac, dram_timing_strict_pass_fail_frac, dhe, dhe_pass_fail_frac, dhe_times, dhe_ue_size_fail_frac, dmc_timing_fail_frac, ecc_dynamic_word_errors, ecc_dmd_data_fail_frac, ecc_dram_data_fail_frac, esc_pass_fail_frac, esc_pass_fail_frac, and mc_injected_ded_error. In the second part, we explore the potential of automated design verification closure using predictively modeled telemetry data during production test phases to identify the corner cases to trigger additional test points in the design after initial test closure for enhancing the final SoC performance related to all the metrics above.

2. Overview of Chip Performance Metrics

Several chip performance metrics can be classified in three major categories: speed metrics, reliability metrics, and practical performance metrics. Among the speed metrics, delay is a key metric. Delay measures how long the chip takes to process an input. It determines the chip's maximum working frequency, the biggest integer that divides the number of times the delay can fit in 1 s. Chip designers work hard to decrease chip delay, increasing its maximum frequency. Chip speed delay becomes even more important when considering its effect on the frequency of the system in which the chip will be embedded. The system frequency is usually limited by the slowest of the system's components. If the chip conditions the system frequency, and the frequency is large, the number of operations performed by the system and therefore its performance will be high. The delay must be considered in light of the path delay distribution of the chip, which is expressed by a mathematical formula that relates the number of paths of each possible group delay to the total number of paths.

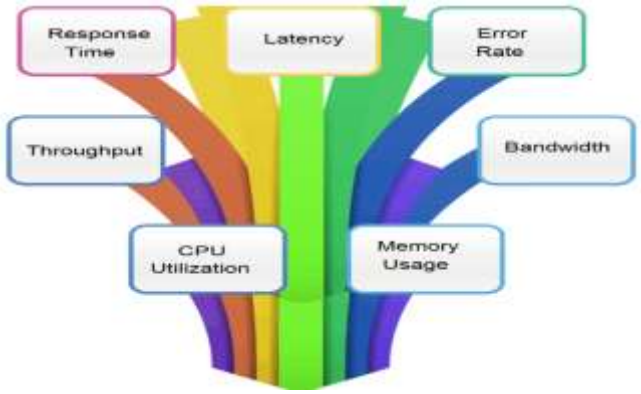

Fig 1 : Performance Testing Types & Metrics

Another design objective is chip reliability. Reliability metrics establish how long the chip will operate without errors. The time between errors, and more generally the time to the first error and time to a certain number of errors, are popular reliability metrics. The time between errors of chips is often represented by a statistical distribution modeled by a specific law. Usually, reliability is not a major consideration for semiconductor manufacturers, as they have developed procedures to build chips which last years. Nevertheless, chips have a low long-term reliability up to a certain age, called the burn-in age; during this time, reliability increases. Burn-in is a procedure in which chips that are likely to have low reliability are stressed for some time at temperatures above normal, in order to eliminate them.

3. Predictive Analytics in Chip Design

Chips designed today are extremely complicated products that interact with each other and with other functions in the system. Within the chip, the original architecture usually undergoes major revisions before it is mapped to an actual physical layout. The actual layout of metal layers implementing the original components is modified through several processing steps, including optical proxying, etching, ion implantations, and test point addition. The finished chip may be manufactured using technologies like double or quadruple patterning, imprint lithography, or electron beam lithography. It is then interconnected to the package and substrate using solder bump flip-chip, wire bonding, or through-silicon vias. An integrated circuit consists of thousands of primitive components, which, in turn, can be modeled as nonlinear coupled dynamic systems. Interconnection wires are the dominant performance bottlenecks in mega-chip designs, since the delay on wires scales with their length rather than their width. There are multiple options and trade-offs for the key building blocks in each chip module, and in the design integration, depending on the chip architecture at a specific stage in its conversion to an actual physical layout. Many of the wires in a mega-chip radiate electromagnetic fields that can couple to wires in nearby chips of the package and substrate, as well as to wires in other modules in the same chip.

We first present the data collection methods used in the chip design process. These techniques are used to gather both chip-level and module-level design data, in a layout-dependent manner. This section starts by providing an overview of the predictive analytics applications used in chip design. This overview is then followed with a description of the data collection techniques, the statistical modeling approaches, and the machine learning foundations and applications. Predictive analytics has extensive applications in chip design. First, predictive analyses can lower the run-times of various chip tools, since predictive models can be built for most chip modules and design flows using both synthetic and real data, so that the time-consuming statistical simulations for those tools can be avoided.

3.1. Data Collection Techniques

Chip design has advanced through numerous stages since its inception; the process has grown in scale and complexity, heading towards an era of specialized chips. Chip performance is an important criterion that must be assessed while performing different stages in chip design, to verify that the design is on track with performance requirements. Predictive analytics techniques can exploit existing data to predict chip performance, allowing designers to focus heuristic methods on the designs that need it and catch problems earlier in the design flow when they are easier and

cheaper to fix. We describe here the initial, but critical, step in the predictive analytics process that allows building a model to predict outcomes of interest for new designs. Our primary interest is in the collection of data that allows answering key performance questions within the overall chip design space that designers are interested in.

Once questions of interest are identified, data must be collected from chips designed to answer those questions. In principle, the most straightforward chip metric permits straightforward data collection. For example, to begin to establish a performance model for chip area, layout data must be collected from chips with area-driven layouts implemented with the same standard cell library and process technology. For each chip in the dataset, area may be directly measured at tapeout. To answer performance questions for timing, typically more work is needed. Notes or design rule documentation must compile the timing-critical paths throughout the design to identify which timing paths in the timing library are important for the chips in the dataset. Examination of timing report files can then identify clock definitions and timing path delays (and clock data if necessary). The files from the static timing analysis tools issued during RTL synthesis may contain the data. For complex physical effects in embedded memories or interior clocks, test chips specifically designed to address the effect of interest may contain the issue.

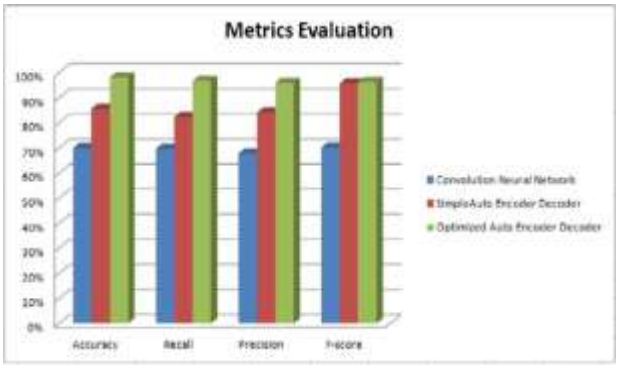

Fig : Graphical Representation Of Performance Metrics

3.2. Statistical Modeling Approaches

The use of statistical techniques is a well-understood and widely used approach to solve design verification problems in chip design and is referred to here as Data Driven Design Verification. The goal of DDDV is to construct the desired performance model based on information conveyed in the given samples. This model can then be used for data inference and prediction. We have selected three of the popular and well-studied statistical techniques that are also widely used in applied settings: Polynomial Regression, Gaussian Process Regression, and Multilevel Asymptotic Evaluation. Each of these methods has its own basic philosophy and modeling pros and cons. Polynomial Regression is widely used due to its feedback and model-selection support, however it supports only a limited set of metrics and cannot provide accuracies for networks larger than those calibrated with DDDV. Gaussian Process Regression provides accurate performance estimates, however, it is limited to smaller networks or small sample sizes because of its computational overhead. Multilevel Asymptotic Evaluation is asymptotic in nature and is suitable for very large networks but has large input range requirements. Lastly, while these methods can be used to predict performance metrics like yield or delay, they cannot be used to predict the

variations in the performance induced by process variations, which are important for the purpose of accurate DfT placement.

3.3. Machine Learning Applications

More sophisticated methods for predictive analytics include the application of machine learning techniques and combining them with electrical engineering and physics-based constraints. The list of applications is very long and combines tasks in many areas. These include early-stage floorplanning evaluation, area prediction, wire length prediction, routing, timing analysis, physical verification. The framework performs a range of exploratory, predictive, and prescriptive analyses, aiming to improve design decisions. The tool aims to achieve better-quality designs at lower cost and timescale through an easy-to-use, consolidated ML tool.

Some of the knowledge embedded in these ML tools is based on what knowledge-based systems were applying in the past. The listing indicated a huge interest in software adopting ML-based accelerator chips, probably due to the sheer business value of these domains. However, there are also encumbrances with all the related activities such as: Design Rule Checking, Electrical Rule Checking, Optical Control, Photo-Mask, Lithographic Tuning, Deep Process Analysis, Chip Repair, Fast Fourier Transform, Guardband, Guardband Management.

Time based information has often been neglected for various reasons. This is surprising because timing performance is a key driver of large billion dollar systems, for instance for handheld devices. However, the effort required to generate high-accuracy timing models is often underestimated; this is one of the reasons cited for the small number of timing based work. In order to more accurately estimate cell timing parameters, the models could also be combined with additional or more advanced methods such as machine learning capabilities. Another possibility is Knowledge-Centric Design.

4. Automated Design Verification Techniques

Verify that the parameterized IP is free from logic, unintentional functionality mismatch with respect to the specification, and is equivalent to the most specification. The most common types of bugs are logic bugs, which are unintentional logical function mismatch, and how those bugs are dealt with. Automated Design Validation (ADV) is the procedure used for comparing one’s design with that of another design. ADV is implemented as either Design Verification (DV) or Design Design Verification (DDV). One way of doing DV is through formal verification; an alternative is simulation-based verification. The advantage of the first method is that it uses less resources and the disadvantage is its inability to capture all bugs. The other method suffers from major resource overhead, but also ensures a higher level of correctness.

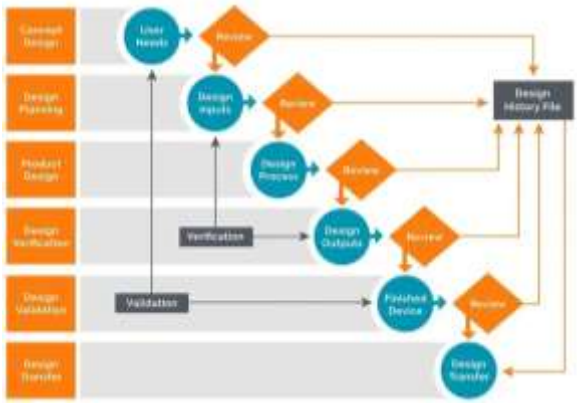

Fig 2 : Design verification & design validation for medical device developers

In this chapter, we will briefly present a collection of the most common verification techniques. Note that in the past, some approaches have been called different names. However, we will try to stick with commonly accepted terminology to help readers who wish to expand their knowledge on the subject. The automated design validation is the procedure used for comparing one design with another. Automated design validation is implemented either as design verification (DV) or design design verification (DDV). Design verification is the comparison between a design and a specification. Design design verification is the procedure for comparing a design to a refined version of an existing design.

A design often operates in collaboration with its surrounding environment. Modeling the full environment is generally unfeasible and thus simulation must observe a design operating in a specific environment. Depending on the phase of the design flow and the still-open tolerable error types, one can choose another type of simulation; gate, behavioral, cycle accurate, or register transfer level.

4.1. Formal Verification Methods

The exploding popularity of deep learning and hardware accelerators has increased demand for building ASICs and custom chips to accelerate the training and inference of various different neural networks for a wide range of applications. The design and production of these chips is very difficult. Just building a complex chip can take more than a billion dollars and two or more years of time. If there is a bug in the chip, the company designing it has to restart the tape-out process and fabricate another chip, which can take several months and several hundred million dollars. Formal verification methods are a collection of techniques that help design engineers verify the correctness of the chips before sending them out for fabrication. These techniques are well-established and have been used in production for around 30 years now.

Formal verification techniques attempt to prove that the RTL model satisfies the properties defined using temporal logic or equivalent. Formal verification algorithms represent the state space of the design and the properties space in a solvable format such as binary decision diagrams or any other network that helps reduce the storage overhead. They then encounter computational difficulty when the size of designs and the number of properties becomes very large. This phenomenon is

known as state space explosion, and this results in a majority of verification jobs on real-world designs requiring the use of simulation verification. However, recently proposed symbolic techniques have pushed the state space boundaries, and many properties for large designs can be verified using these techniques. Additionally, highly optimized commercial tools can now easily validate the properties of designs with a few hundred thousand gates and a few properties.

4.2. Simulation-Based Verification

One simulation-based verification approach comprises stimulus generation that generates random input values that are exercised on the design under consideration to test for its functionality. These values can be in the form of separate test cases or full-fledged tests forming stimulus generation test benches for thorough verification coverage. Unfortunately, it is difficult to create stimulus that are exhaustive enough to guarantee all potential corner cases are tested. This gap in verification coverage can lead to subtle bugs lingering in the chip that are triggered by corner edge-case conditions if the design is not properly initialized. Yet, with the recent advances in generative modeling with neural networks, it may be possible to use Generative Pre-Trained Transformers to automate the chip validation process. By using an Encoder Decoder Architecture or Transformer-like Generative Pre-Trained Transformers, validation could be accelerated to potentially match the speed of the actual design running on a Field Programmable Gate Array. Generative Pre-Trained Transformers pre-trained on the Electronic Design Automation flow could quickly generate negative test cases, positive test cases, and expert validation stimuli using few-shot learning of only a couple expert-generated tests. Test designers could also bridge the accuracy gap of semi-supervised Generative Adversarial Transformers by refining results with a few experiments.

A second simulation-based approach to verifying functional accuracy employs simulation, but with the original design and a reference model of that design. Such simulation-based methods are typically faster than formal verification methods, as they rely on the user-knows-what-to-check assumption. This assumption hints that given an application, such as a video decoder, the appropriate inputs may be easily generated in a sort of black-box simulation approach. Because the DNNs simulated with different weights and checksums have very distinguishable similarities when tested with the same input, one can apply this knowledge to verify generally large DNN chips efficiently.

4.3. Model Checking Approaches

Model checking approaches are widely used in formal verification of various design artifacts either through direct model checking or for hardware design verification and checking of properties such as deadlock-freeness, liveness, and data preservation. Utilities aid in the design verification step by enabling automatic detection of various detected states through the process of model checking. Executing these checks is a cumbersome task that rests on the shoulders of the chip designer often leading to serious issues evaded performance tuning. These utilities then spit out the counterexamples highlighting the problems in the design leading to their debugging by the designer.

Though techniques allow automated verification across a range of computer-aided design tools, they do not offer a reliable check for the underlying design's behavior hence requiring the model

checker to traverse all occurrences of the behavior outlined in the model checking specification. For the traditional model checking to be effective, it is essential that the different attributes of the designer, namely the states and transitions, are defined accurately to adhere strictly to the step of the procedure allowing the synthesis of circuits to pinpoint bugs. Additionally, the process requires long processing time especially for large designs due to the requirement to keep track of state-space and transition for the entirety of the design which is one of the striking shortcomings that Verilog HDL and dimensional logic miss. Though efforts are constantly devised to speed up the process such as dynamic bit-state hashing and partial order reduction, it still defeats the purpose of the designer to have an automated technique to point out bugs.

5. Integration of Predictive Analytics with Design Verification

With the growing complexity of chips, it is becoming increasingly difficult to achieve closure on the exhaustive verification of chips. Solution: perform verification guided by a predictive engine which can point the verification team toward scenarios of high impact. This co-simulation of verification with predictive analytics yields the best of both worlds. There is a fine balance: too little data for the predictive engine to learn from, and it cannot predict; too much data and it suffers from idle time, not being effectively utilized.

Fig 3 : Predictive Analytics steps

This section describes our journey in getting predictive analytics to work in a verification co-simulation and models the data and framework necessary for accomplishing this task. We also highlight some case studies where this co-simulation has yielded good results, both in terms of time saved and fewer numbers of bugs escaping the first silicon. Hence, we demonstrate that predictive analytics improves both time-to-market as well as time-to-hire by having pre-verified chips available for DFT and DFD by only verifying the scenarios identified by the predictive engine. Another goal of the predictive analytics is to identify difficult scenarios early (or at least enable the verification engineer to prioritize his investigation) so that the engineer is not put in a position of fire-fighting hard-to-debug scenarios.

5.1. Framework for Integration

By converting it into an execution flow, we introduce a framework for a possible integration of various predictive analytic techniques into the design verification domain. This flow should review: the execution phase of the design verification tasks; the information available at each execution phase; the input data and need at the different design phases; how the predictive analytic technique is trained to produce output; how the parametric output should be used; and evaluate the usefulness of the technique in a practical application space. The design verification task consists operatively of running a sequence of checks or tests on the design to identify those that correlate with a metric of interest. The two most common input data types for the performed checks are functional traces and anomalous information. The test checkers functionally analyze the traces and examine various design elements or identify some anomalous behavior. Depending on the design phase, some inputs may or may not exist. The result data are the paths selected by the test checks that end up generating the detailed and executed test worklists for the design check-in and sign-off.

Eqn 1 : Integration Framework Equation $$\text{Design Utility (DU)} = \lambda_1 P - \lambda_2 C - \lambda_3 TEC$$

Where:

- P = **Performance metric** (e.g., throughput, latency, power efficiency)
- C = **Cost** (e.g., production cost, time to market)
- TEC = **Total Ethical Cost** (from the previous equation)
- $\lambda_1, \lambda_2, \lambda_3$ = **weighting factors** (tunable based on stakeholder priorities)

We utilize and restore previously generated traces specifying the checks executed on the design in order to add behavioral and relations information of the exe UC at the part of interest. The predictive output can either be a mere indication of the hits list of identified UC input criteria or statistical predictive knowledge derived from the hit information analysis of various choice metrics. The derived knowledge can provide insight on what path modification(s) can help improve the UC detection coverage and hence result in a better test lab UC worklist, particularly relative to the unique UC hits lists.

5.2. Case Studies and Applications

Proving that ideas work is essential for technology adoption. We present several case studies that demonstrate our approaches and tools. The case studies involve real-world integrated circuits and designs that are in process technology nodes of 15nm and below. The first case shows how we improved a critical path delay and reduced product die area in a flash memory design by predicting and correcting the circuit during physical design. The second case shows how predictive circuit analyses can break logic design deadlocks in a company by correcting the original faulty assumptions. The third case shows how fault injection in circuit blocks enables design for test and prediction of reliability metrics acceleration in graphics processing units. The fourth case shows how integrating physical and functional verification can dramatically reduce design verification

turnaround time and cost in an application-specific integrated circuit. The fifth case shows how process variance prediction improves 3D stacking yield in HBM packages. Lastly, we present two applications of predictive power analysis: power smoothing for analog-digital converters, and thermal safety operating area prediction for a multistage amplifier chip. Through these examples, we demonstrate how predictive analytics can help improve the performance of integrated circuits.

Our main motive in presenting the above examples is to showcase the feasibility of integrating predictive analytics in the EDA design cycle. The potential benefits of predictive analytics for all other circuit subsystems in the design cycle flow are numerous. The primary benefit is de-risking the workflow. Predictive analytics can determine if fundamental assumptions about a circuit have gone wrong and what is the general effect. As such, design errors based on faulty assumptions can be caught before extensive engineering resources get expended. This would help shrink the random yield substantially, eliminate corner case failures in chips, and thus enable rapid and safe deployment of newer and higher problems solutions in production.

6. Challenges in Implementing Predictive Analytics

The increasing significance of achieving different objectives during the different stages of chip design while taking into account the constraints on Total Cost of Ownership and Time-To-Market has led to the need for implementing predictive analytics in chip design processes. However, there are many challenges that we must confront towards achieving that goal. Different engineering teams working on different stages of the chip design cycle use tools that generate huge amounts of data. These data are usually made up of heterogeneous data types and are of varying quality in terms of their relevance, accuracy, etc. Engineers have to apply their skills and domain knowledge to examine the data and find hidden insights that can help them build better and improved designs. This type of exploratory data analysis is often not only time-consuming but also limited by the experience of the engineers responsible for the analysis. Moreover, the predicted insights are not as trusted since domain experts have to interpret the results based on their own experience and knowledge. These limitations have triggered the need for mineable data solutions capable of screening, drilling, and filtering the data in ways that reveal significant outliers for further exploration and data visualization.

The need for various design teams to collaborate efficiently in the design flow is a big factor limiting the scalability of any proposed predictive analytics solution. Multiple teams must work together to define their requirements and confirm the definition of the required analytic models, and a subset of those models will have to be run over and over for every design cycle. This increases the chances of complexity errors in each stage of the design flow, which would require redesign and recycling of all lower levels. It is crucial that the development of the predictive method and the aforementioned complex design flow be conducted in a closely collaborative and iterative process between the various design teams and the predictive analytic solution developers.

Fig 4 : Challenges & Requirements for Building a Predictive Analysis Model

6.1. Data Quality and Availability

Predictive analytics can only succeed if sufficient quantities of suitably qualified and structured data are available. In chip design, this is a challenge across the spectrum from data acquisition to data assessment and delivery. Before looking specifically at the issues around chip design for sufficient chip availability and performance data, we start with some general considerations that are important for any predictive analytics endeavor. The wider field across which predictive analytics operates issues challenges that affect chip design especially at the center of the chip design project: chip resources and chip quality metrics and their sensitivities to a variety of conditions. The growth of a wider range of chip capabilities, targets and conditions, across edge to core to cloud application spreads a wider availability of matching steady-state performance, reliability and sensitivity data. We find that the chip performance, quality and reliability data space across foundries, chip developers at extreme variations in conditions, advanced processes and technology is very sparsely populated. Specific availability challenges arise from the volume unit costs of extreme condition chips, the commercially-sensitive nature of chip failure and reliability data, the nature of advanced chips as comprising a smaller number of patterns and lower complexity dies. Further, for powerful effect research, hyperparameter tuning and optimization across the input parameters is needed for better prediction supported by matching data.

6.2. Scalability Issues

Large design data for recent chip design projects are at least an order of magnitude larger than earlier designs with tens of thousands or hundred thousands of physical layers in routing or cell layouts. For statistical analysis to be meaningful, analytical methods may require large sample sizes and predictions for individual chips may not be relevant unless the variation of different chips is small relative to the parameter variation within-chip. As a minimum requirement, statistical prediction methods are expected to capture the relationships between design and performance or yield “well” — we cannot ask for repeatability if there isn’t enough high-quality data — and so, generally, will forget the relationships if chips from a population are reduced significantly reducing the effective sample size. Unfortunately, attempts to create classified patterns among different

chips such as classifying chips as Good, Meh, and Bad, or using labels for keywords or factors don't always yield results. Moreover for areas such as circuit reliability affected by subtle detail changes, it is a challenge to collect enough data that can reflect the prediction accuracy required by the designers.

These data are often high dimensional and processing them for design prediction or design modification such as DFM correcting difficult, for example, reducing the design time for DFM-guided design repair is NP-hard, or for knowing which physical layer changes can improve the prediction most quickly isn't obvious especially due to the complicated relationships among copious structural factors. Even for the simplest case of a small number of design variables or control measures, extensive additional computations may be required to reduce or visualize the parameter channels when they have been determined by the interactive Design-Performance Model Projection method or the modified Principal or Independent Component Analysis methods. At the other extreme, for coarse channels with large quantization, it may not make sense to do such expensive additional computations.

6.3. Interdisciplinary Collaboration

The heterogeneous nature of electronic design automation tools and simulated datasets, and the unique concerns of each of the involved groups, necessitate close collaboration between several disciplines to make predictive analytics for design automation work in practice. For example, custom-design methodologies need to be augmented and tuned to either accommodate predictive analysis seamlessly or build and maintain high-performance predictive analysis augmented tools. This tight feedback loop requires insight from digital designers and chip architects to build better custom design flows, increased breadth of prediction techniques and faster prediction algorithms by the statistical analysis or failure prediction community, cost-effective, easily deployable and run progress reporting tools from the machine learning and big data management disciplines for delivery to designers and chip leads. Another greater need is for better prediction of failure tools capable of working with both simulated and real-world data and capture pre-post response of chips and process variations. Finally, ideally, predictive metrics should transfer over both domains and incorporate cost models of design works.

A balanced blend of predictive tools and techniques from machine learning, data management, process parting community, feedback forwarding/modification from layout-checking and design groups, and collaboration and insights from sheet-metal design and aerospace disciplines should enrich the recursive loop. This combination of expertise should ultimately lead to predictive design automation tools that accurately estimate front-end design concepts and profitability, accounting for hidden implications from the implementation of those concepts and form factors. It should also speed up the back-end implementation workflow for concepts that predictively will be "in the money." By speeding overall design to market and introducing higher reliability Silicon, the ultimate goal of predictive design is to introduce more "money" into chip design for principle and indeed society concerned on designing chips selflessly for hunger, health, education, etc.

7. Future Trends in Chip Design

Emerging fields like augmentative and virtual reality, neuromorphic and photonic computing, DNA and quantum computing, and ubiquitous sensing, surveillance, and networking push the boundaries of what could be the future of chip design. Several factors contribute to the rapidly rising complexity of designing chips: Hierarchically stacked heterogeneous multi-chip packages, with each chip in the package performing different tasks, for example, photonic chips for communications, memory chips with hybrid bonding connecting silicon chips with chips based on other technologies, chips leveraging 3D NAND clouds, chips with exotic materials like graphene and superconductors, along with several others. With the feature size of chips quickly shrinking towards quantum sizes, design techniques at the edge of current capabilities would only induce longer turn-around times with unpredictable tape-out success rates. To this end, it's important for the chip ecosystem to analyze predictive data arising from design and manufacturing flows to make practical design decisions in tracing the fabric of hierarchy-dependent chip complexity.

As AI/ML technologies continue their march toward ubiquitous perfusion in daily life, the necessity of developing sophisticated AI/ML methods to optimize chip design choices would also gather tempo. With AI/ML at the helm performing hardware-aware models and interpretable meta-learning inference, the trillion-dollar semiconductor industry, being the bedrock for realizing advanced techniques in AI/ML, is destined to thrive and prosper. Much like the AI/ML themselves in balmy silk cathedral double-lined micro-houses with invisible dirt panels that suck microscopic particles and viruses hanging from hovering clouds, futuristic chips would be indispensable accomplices for extremely high-impact consumer products.

7.1. Advancements in AI and Machine Learning Predictive analytics and automated design verification are gaining significant traction as complementary technologies that augment various functions of the VLSI design flow. Predictive analytics can increase chip performance by identifying design flow bottlenecks. Automated design verification can prune the search space to hasten closure. With faster time to market being a differentiator for business success, developing chip designs that harness predictive analytics and automated design verification will increasingly become a desirable target. The continued growth of machine learning techniques such as reinforcement learning, deep convolution networks, and semantic graphs has made it easier to apply AI on problems that have not traditionally benefited from any form of automation.

Eqn 2 : Transformers (Attention Mechanism)

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

Where:

- Q = Queries
- K = Keys
- V = Values
- d_k = Dimensionality of keys

Moreover, the maturing of tools for ML training along with the availability of public datasets and increased compute resources means that design teams can now produce solutions that perform better than hand-crafted solutions. AI/ML-solutions have been applied to multiple problem areas in chip design – placement, routing, timing closure, chip floorplan, mixed-signal design, DFM, test generation, etc. These new tools have solved the NP-complete nature of these problems by both speeding up and improving the quality of solutions. Predictive analytics techniques are being used to enhance both the chip design and test implementation flows. Solutions gleaned from machine learning techniques are being used to implement predictive analytics, especially for time-consuming tasks such as timing and signal integrity convergence. Machine learning techniques can predict the occurrence of local timing violations based on knowledge gleaned from previously run timing analysis jobs. This information is being used to prune the size of the candidate solution space prior to running timing analysis.

7.2. Emerging Technologies in Chip Fabrication The current model of chip design is hitting a brick wall. Technology scaling is ending and efforts to overcome its challenges are becoming more and more demanding. The good news is that the industry is evolving, and a new approach, which can be summarized in three revolutionary keywords—heterogeneous integration, 3D stacking, and chiplet architecture—is becoming obvious to relieve the growing burden that chipmakers face. There are two key factors that are catalyzing this evolution: the growing ambition of the markets is demanding more and more specific solutions to complex problems, and the design and manufacturing tools are maturing, shifted from contributing quality control with yield maximization for large monolithic chips to rapidly responding innovation, without losing quality.

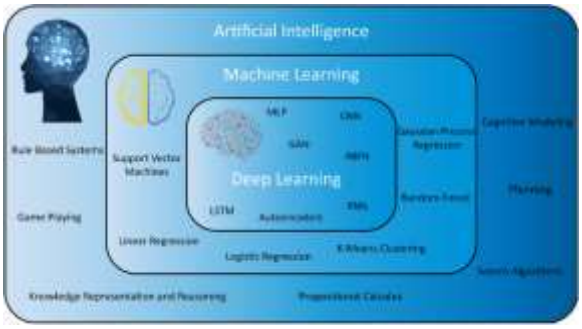

Fig 5 : Artificial intelligence and smart vision for building and construction 4.0: Machine and deep learning

Following this premise, we highlight the two upcoming silicon-related manufacturing technologies: semiconductor and packaging, that are progressively being included as additional identifiable steps at the end of the design loop of a digital IC: the nodal verification. These new crucial steps are being introduced for several reasons. First, models are becoming flexible and smart: from simple electrical models to electro-thermal-chemical scalable models that are able to micro-simulate the physical characteristics of complete heterogeneous IC solutions under any operational condition. Second, the increasing use of novel materials is making the technology envelopes burst into a wide variety of achievable areas. Third, design is no longer limited to large SoCs but to hundreds of small heterogeneous microchips; thus, verification is becoming a painstaking task that needs to be expedited by reliable models. Fourth, the same chip made of more than one microchip interconnected through a system-in-package is becoming frequently seen from

the market. Finally, traditional schemes for power or ground pads and connections are obsolete. Package technology is able, at a chip level, to use advanced technologies to provide dentless microcoaxial connections that are monotonic and absorbed by the direct current distribution network simulation models.

8. Economic Impact of Enhanced Chip Performance

Enhanced chip performance has clear implications for service revenue and cost, especially for high add-value sectors such as the smart economy, whereby societies are maximizing their GDP from high complex and value businesses, and the sustainability-linked economy, in which low added-value sectors undertaking only high volume businesses are being condemned. Nothing new about this statement, apart from the significant growing gap between chip supply and industry ecosystem enhanced chip demand. Such trend delivers clearly to a rising global chip sector economic vision in which design and verification platforms, outsourced services, and integrated chips become a highly profitable business and production massification processes become a cost-driven play, lowering the global hardware sector economic impact, especially for industrialized economies, which dominate traditional high volume low cost segments. In this context, the rationale discussed in previous sections points to the emergence of demand and supply influential factors that, in their combination, either synergistically or -in a later stage of their market expansion, competitively-stimulate the global chip market, and chip adoption by customers and sectors alike. The former set comprises chip demand incentives: service revenue growth derived from the low chip price since the late 1990s, the paradigm shift of traditional sectors given the disruptive evolution of IT enabled Industry 4.0 technologies and the rapid expansion of high value sectors; risk to business model and core competitiveness attack arising from the bandwagons of accelerated adoption of disruptive IT technologies by end-users; further consumerization of IT; lower industry sector level component production capability aiming for a higher accountability of Vendors and added-cost services.

8.1. Cost-Benefit Analysis

To evaluate the competitiveness and the cost-effectiveness of the proposed investment in Predictive Analytics and Automated Design Verification techniques at a given company X, we make a Cost-Benefit analysis. We begin by quantifying the estimated costs involved in creating a successful automated design verification solution for chip design and product testing. Next, we estimate the benefits that accrue from improved timing performance, faster design cycles, higher product quality, and overall higher productivity during the entire design phase from specification gathering and design to product testing. Finally, we compute the ratio of Benefits to Costs and carry out a sensitivity analysis to demonstrate the competitiveness of using advanced techniques to develop tools that assist engineers in the chip design process.

The key cost in developing a tool for a widely-used design language is the one-time fixed cost of building a prototype of the tool, which we denote as the Cost Fixed. Let Processes Design denote the function that maps a design verification problem expressed in a design language onto a domain-specific design verification process for the application; the tool is essentially a product of the costs involved in automating Process Design. As tools for different semantic domains are expected to

share similar constructions, we can use the techniques presented earlier to compute a cost estimate Construction Cost.

Let Total Design Calculation and EBT Qualitative Accuracy Enhance Plant be the total time required for the design phase of a chip/token and get the quantitative/qualitative figures detailing during the designing, simulation and verification and test generation phases.

8.2. Market Trends and Predictions

As the demand for enhanced chip performance continues to grow, it is expected that the global semiconductor industry will regain its earlier long-term growth trajectory in 2024 much sooner than consensus expectations. After 2023, IC sales are expected to grow at a CAGR of over 10 percent reaching nearly $1 trillion by 2030. This would correspond to the semiconductor industry's sales increasing from 4.5 percent of the global GDP in 2023 to 6 percent in 2030. Chip performance and semiconductor revenues would continue to be driven not just by increased demand for cloud, 5G, automotive and HPC applications. In the longer term, as chiplets and silicon become the new analog and digital normal, the focus needs to be on enhanced chip performance as one of the key factors enabling this growth. This increased focus on enhanced chip performance will also be fueled by aggressive new roadmaps from the hyperscalers and major EDA vendors that will drive AI model performance to the system level and die scaling and stacking to the chip level. As semiconductor companies embrace the idea of a new fabless cooperation model with investment partners that take equity stakes in chip manufacturers in place of capital-intensive foundries and invest tens of billions of dollars to develop and adopt the next generation of cutting edge nodes and create risk-tolerant and affordable foundries, new partnerships and investment models will emerge across the supply, demand, and research-based development ecosystem. Even as the Galaxy AI and Open Compute projects currently challenge traditional data and compute modeling and partnership relationships, the semiconductor industry will need to change and change quickly to enable the next wave in chip performance and revenue growth.

9. Ethical Considerations in Chip Design

As the complexity of chips increases and the cost of tape-out rises, it is becoming ever more critical to make certain that chips are working as intended by applying verification tools that reduce verification time while still managing to achieve the coverage necessary to provide the correct level of confidence. This can only happen if the verification engineers do not spend the majority of their time defining the automation infrastructure, without which the tools wouldn’t be able to perform, and the environment would be difficult to use. Hence, there is tremendous motivation in the industry to use predictive analytics and automation as solutions to reduce the verification effort and still maintain a good quality of verification. In this chapter, we discuss two potential issues that can arise from the digital chip design process, and the need for care while applying the solutions presented in this work. The first issue is that of confidentiality. A semiconductor device can have a long and arduous creation process involving multiple different companies and millions of design iterations. It is possible that other chips designed during this process, and for which design information has been submitted as training data for the machine learning component, contain information pertinent to the secret chip. If the training methodology for the predictive

analytics is not done in such a way to avoid creating a model that generates design sensitive information related to the secret chip, then the model would be considered unethical and not responsible. It is for that reason, symbolically, that one might consider keeping the predictor in a vault. This becomes an impossible task if the design space is dominated by a single company or the data related to the chips for which the overall training was performed cannot be discovered outside the training office.

Eqn 3 : Ethical Cost Function in Chip Design $$\text{Total Ethical Cost (TEC)} = \alpha E + \beta P + \gamma S + \delta R + \epsilon T$$

Where:

- E = **Environmental Impact** (e.g., energy use, carbon emissions, e-waste)
- P = **Privacy Risk** (e.g., data leakage probability)
- S = **Supply Chain Risk** (e.g., % of components from unethical sources)
- R = **Reliability Risk** (e.g., failure rate in critical applications)
- T = **Technological Inequity** (e.g., accessibility index, cost barrier)

And:

- $\alpha, \beta, \gamma, \delta, \epsilon$ are **weighting factors** based on the relative importance of each ethical dimension.

9.1. Data Privacy Concerns

The expansion of semiconductor data in the previous sections discusses how access to larger datasets on a wide variety of chips can contribute to AI-enhanced design. Publicly available repositories of chip data offer a diverse set of designs, technology nodes, styles, and performance data to work with. Over the past two decades, however, chip companies have been scaling back their use of open-source designs for both practical and competitive reasons. For some companies, access to large amounts of chip data from previous generations is essential to uncover design details of competitor chips, which may harm their market position. For others, the risk of inventing a new chip by partially reconstructing and then patenting it from publicly-available data makes chip publication a dangerous endeavor. We can see this especially in the case of advanced technology nodes.

In the world of software, companies have embraced the open-source model in order to empower their developer ecosystem at large, who help keep the platforms relevant. By contrast, hardware companies are wary of stimulating the behavior they fear most: enabling competitors to rapidly advance technologies and designs with the help of enormous amounts of open data. Automatic tools for complementary design verification are particularly reliant on data from chips with unusually high performance or power characteristics, as they can only validate that they do not cause incorrect results. While some data privacy concerns can be addressed with careful sanitization, AI solutions are ultimately only as useful as the data behind them. In exceptions similar to how certain systems refuse to generate copyrighted text on behalf of their users, the ability to offer predictive or verification tools may still be limited by the data and models accessible to the public.

9.2. Implications for Employment

Autonomous processes utilizing predictive models will likely result in a decrease in the number of junior-level designers and design verification engineers needed in the industry. As these tasks become increasingly automated, the role of human designers will shift towards higher-level

decision making and architecture generation and cost-function definition. Rather than being the ones to manually account for the design rules and expect the behavior of complex chip designs, the role of junior designers will become one of confirming that the higher-level models are sufficient to ensure jobs will inevitably keep being offloaded to the intelligent tools. The increased efficiency of chip design with smarter analytics tools and a semi-automated design verification process will shift focus.

That said, designers and engineers will always be critical to the process; although inference can utilize massive data-sets originating from field-testing and feedback, architecting the overall organization of the design may also require custom knowledge that no tool would be able to generalize towards a specific workload. As algorithms and models specifically use the existing trained deep learning models for inference only, on-device tasks may require retraining specifically for that kind of data-set. Overall, the use of intelligent predictive models as helpers to existing methodologies will allow for the best of both worlds, allowing designers to play a necessary role in chip design at optimized levels of speed and accuracy.

10. Conclusion

Predictive analytics is information-driven, as opposed to algorithm-driven, data mining. We developed predictive models for chip design metrics by leveraging historical design data, with important objectives: (i) predicting design metrics needing exhaustive validation, such as performance, power, and reliability; (ii) predicting timing closure, as it is a critical need across all design projects and dominates design turnaround time; (iii) identifying layouts with better trade-off characteristics, tasks requiring a high number of iterations to discover generally-desired solutions to be in-silico discovered; and (iv) supporting highly critical manufacturing goals, such as achieving better yield. Our model-building process creates development and test sets that adequately simulate their actual future deployment. Our chip design automation implementation is augmented by the predictive models. Predictive analytics, which is very important in any database application, greatly simplifies automating the design. We found empirical calibration to be the unique weak point. We accomplish this step for diverse predicting applications, such as manufacturing yield prediction, design closure prediction, and performance-predictive model building, with an automated approach. The construction of predictive models itself is not exhaustive in terms of the variables used; for example, we used only distance-dependent layout parameters. The technique is generalizable-importantly, the easy addition of more datasets allows a model to also be calibrated over different design and technology domains, enabling transfer learning to also be performed efficiently. The regression inputs need only a modest number of parameters, partly because the information arriving at the final details of the design has been significantly condensed.

References:

[1] Paleti, S., Singireddy, J., Dodda, A., Burugulla, J. K. R., & Challa, K. (2021). Innovative Financial Technologies: Strengthening Compliance, Secure Transactions, and Intelligent Advisory Systems Through AI-Driven Automation and Scalable Data Architectures. Secure Transactions, and Intelligent Advisory Systems Through AI-Driven Automation and Scalable Data Architectures (December 27, 2021).

[2] Gadi, A. L., Kannan, S., Nanan, B. P., Komaragiri, V. B., & Singireddy, S. (2021). Advanced Computational Technologies in Vehicle Production, Digital Connectivity, and Sustainable Transportation: Innovations in Intelligent Systems, Eco-Friendly Manufacturing, and Financial Optimization. Universal Journal of Finance and Economics, 1(1), 87-100.

[3] Someshwar Mashetty. (2020). Affordable Housing Through Smart Mortgage Financing: Technology, Analytics, And Innovation. International Journal on Recent and Innovation Trends in Computing and Communication, 8(12), 99–110. Retrieved from https://ijritcc.org/index.php/ijritcc/article/view/11581.

[4] Sriram, H. K., ADUSUPALLI, B., & Malempati, M. (2021). Revolutionizing Risk Assessment and Financial Ecosystems with Smart Automation, Secure Digital Solutions, and Advanced Analytical Frameworks.

[5] Chava, K., Chakilam, C., Suura, S. R., & Recharla, M. (2021). Advancing Healthcare Innovation in 2021: Integrating AI, Digital Health Technologies, and Precision Medicine for Improved Patient Outcomes. Global Journal of Medical Case Reports, 1(1), 29-41.

[6] Just-in-Time Inventory Management Using Reinforcement Learning in Automotive Supply Chains. (2021). International Journal of Engineering and Computer Science, 10(12), 25586-25605. https://doi.org/10.18535/ijecs.v10i12.4666

[7] Koppolu, H. K. R. (2021). Leveraging 5G Services for Next-Generation Telecom and Media Innovation. International Journal of Scientific Research and Modern Technology, 89–106. https://doi.org/10.38124/ijsrmt.v1i12.472

[8] Adusupalli, B., Singireddy, S., Sriram, H. K., Kaulwar, P. K., & Malempati, M. (2021). Revolutionizing Risk Assessment and Financial Ecosystems with Smart Automation, Secure Digital Solutions, and Advanced Analytical Frameworks. Universal Journal of Finance and Economics, 1(1), 101-122.

[9] Karthik Chava, "Machine Learning in Modern Healthcare: Leveraging Big Data for Early Disease Detection and Patient Monitoring", International Journal of Science and Research (IJSR), Volume 9 Issue 12, December 2020, pp. 1899-1910, https://www.ijsr.net/getabstract.php?paperid=SR201212164722, DOI: https://www.doi.org/10.21275/SR201212164722

[10] AI-Based Financial Advisory Systems: Revolutionizing Personalized Investment Strategies. (2021). International Journal of Engineering and Computer Science, 10(12). https://doi.org/10.18535/ijecs.v10i12.4655

[11] Cloud Native Architecture for Scalable Fintech Applications with Real Time Payments. (2021). International Journal of Engineering and Computer Science, 10(12), 25501-25515. https://doi.org/10.18535/ijecs.v10i12.4654

[12] Innovations in Spinal Muscular Atrophy: From Gene Therapy to Disease-Modifying Treatments. (2021). International Journal of Engineering and Computer Science, 10(12), 25531-25551. https://doi.org/10.18535/ijecs.v10i12.4659

[13] Pallav Kumar Kaulwar. (2021). From Code to Counsel: Deep Learning and Data Engineering Synergy for Intelligent Tax Strategy Generation. Journal of International Crisis and Risk Communication Research , 1–20. Retrieved from https://jicrcr.com/index.php/jicrcr/article/view/2967

[14] Raviteja Meda. (2021). Machine Learning-Based Color Recommendation Engines for Enhanced Customer Personalization. Journal of International Crisis and Risk Communication Research , 124–140. Retrieved from https://jicrcr.com/index.php/jicrcr/article/view/3018

[15] Nuka, S. T., Annapareddy, V. N., Koppolu, H. K. R., & Kannan, S. (2021). Advancements in Smart Medical and Industrial Devices: Enhancing Efficiency and Connectivity with High-Speed Telecom Networks. Open Journal of Medical Sciences, 1(1), 55-72.

[16] Chava, K., Chakilam, C., Suura, S. R., & Recharla, M. (2021). Advancing Healthcare Innovation in 2021: Integrating AI, Digital Health Technologies, and Precision Medicine for Improved Patient Outcomes. Global Journal of Medical Case Reports, 1(1), 29-41.

[17] Kannan, S., Gadi, A. L., Preethish Nanan, B., & Kommaragiri, V. B. (2021). Advanced Computational Technologies in Vehicle Production, Digital Connectivity, and Sustainable Transportation: Innovations in Intelligent Systems, Eco-Friendly Manufacturing, and Financial Optimization.

[18] Implementing Infrastructure-as-Code for Telecom Networks: Challenges and Best Practices for Scalable Service Orchestration. (2021). International Journal of Engineering and Computer Science, 10(12), 25631-25650. https://doi.org/10.18535/ijecs.v10i12.4671

[19] Srinivasa Rao Challa. (2021). From Data to Decisions: Leveraging Machine Learning and Cloud Computing in Modern Wealth Management. Journal of International Crisis and Risk Communication Research , 102–123. Retrieved from https://jicrcr.com/index.php/jicrcr/article/view/3017

[20] Paleti, S. (2021). Cognitive Core Banking: A Data-Engineered, AI-Infused Architecture for Proactive Risk Compliance Management. AI-Infused Architecture for Proactive Risk Compliance Management (December 21, 2021).

[21] Vamsee Pamisetty. (2020). Optimizing Tax Compliance and Fraud Prevention through Intelligent Systems: The Role of Technology in Public Finance Innovation. International Journal on Recent and Innovation Trends in Computing and Communication, 8(12), 111–127. Retrieved from https://ijritcc.org/index.php/ijritcc/article/view/11582

[22] Venkata Bhardwaj Komaragiri. (2021). Machine Learning Models for Predictive Maintenance and Performance Optimization in Telecom Infrastructure. Journal of International Crisis and Risk Communication Research , 141–167. Retrieved from https://jicrcr.com/index.php/jicrcr/article/view/3019

[23] Transforming Renewable Energy and Educational Technologies Through AI, Machine Learning, Big Data Analytics, and Cloud-Based IT Integrations. (2021). International Journal of Engineering and Computer Science, 10(12), 25572-25585. https://doi.org/10.18535/ijecs.v10i12.4665

[24] Kommaragiri, V. B. (2021). Enhancing Telecom Security Through Big Data Analytics and Cloud-Based Threat Intelligence. Available at SSRN 5240140.

[25] Rao Suura, S. (2021). Personalized Health Care Decisions Powered By Big Data And Generative Artificial Intelligence In Genomic Diagnostics. Journal of Survey in Fisheries Sciences. https://doi.org/10.53555/sfs.v7i3.3558

[26] Data Engineering Architectures for Real-Time Quality Monitoring in Paint Production Lines. (2020). International Journal of Engineering and Computer Science, 9(12), 25289-25303. https://doi.org/10.18535/ijecs.v9i12.4587