Instituto Tecnológico y de Estudios Superiores de Monterrey

Estado de México Campus

School of Engineering and Sciences

Doctor of Philosophy in Computer Science

Dissertation proposal

A Hybrid Multi-objective Optimisation Approach to Deep Neuroevolution for Super-Resolution Image Restoration

by

Jesus Leopoldo Llano Garcia

ID A01748867

Atizapán de Zaragoza, Estado de México, June, 2021

Instituto Tecnológico y de Estudios Superiores de Monterrey

Campus Estado de México

The committee members, hereby, recommend that the proposal presented by Jesus Leopoldo Llano Garcia to be accepted to develop the thesis project as a partial requirement for the degree of **Doctor of Philosophy in Computer Science**.

Dr. Raúl Monroy Borja
Tecnológico de Monterrey
Principal Advisor

Dr. Víctor Adrián Sosa Hernández
Tecnológico de Monterrey
Co-Advisor

Dr. Leonardo Chang Fernández
Tecnológico de Monterrey
Committee Member

Dr. Carlos A. Coello Coello
Centro de Investigación y de Estudios Avanzados del IPN
Committee Member

Dr. Miguel Angel Medina Pérez
Tecnológico de Monterrey
Committee Member

Dr. Hugo Terashima Marín
Director of Program in Computer Science
School of Engineering and Sciences

Monterrey, Nuevo León, June, 2021

Contents

List of Figures

List of Tables

Abstract

Super-resolution image restoration poses an inverse problem, focusing on obtaining a high-resolution image from a low-resolution corrupted sample. It has been well-studied in computer vision and image processing, with relevant importance in many fields ranging from security and surveillance to image generation. Contemporary advances focus on using deep neural networks to tackle these problems. However, most current neural network architectures used for super-resolution continue to be developed by human experts through time-consuming and error-prone trial and error.

Nowadays, there is a growing interest in automated neural architecture search methods that allow deep learning to be more accessible to researchers and practitioners. Notwithstanding the success of these methods, their application in real-world problems still poses significant challenges, hindering their practicality. In numerous instances, the synthesized architectures are too complex to be deployed in resource-limited platforms or embedded systems. In contrast, architecture discovery requires copious amounts of processing time and power. To get around these problems, one solution that started to grow in popularity is to see the process of architecture search as a multi-objective problem, considering execution latency, memory footprint, and model performance, among many others. Yet, adding additional objectives to the search increases the complexity of identifying the most efficient architecture.

While other works have studied the application of simple local search algorithms for discovering and improving existing topologies, few have explored the use of both. Our work will combine a local search algorithm with a population-based global search technique, aiming at an increase in the speed at which solutions move toward optimality. We will study the advantages and limitations that the combination of these techniques should instill within the process of neural architecture search.

Thus, this research's goal is the design of a multi-objective neural architecture search hybrid evolutionary algorithm for super-resolution image restoration. Tackling the problem of architecture search from a hybrid multi-objective perspective strives for synthesizing resource-efficient models capable of obtaining a competitive performance in Super-Resolution tasks while enhancing the efficiency of architecture search. This approach should not only facilitate the deployment of deep learning models but also help study the characteristics of architectures that present different trade-offs among performance and complexity, such as the number of parameters, number of floating-point operations, training efficiency, and inference latency.

1 Introduction

Many Deep Neural Networks (DNN) have shown outstanding results attributed to their ability to automate feature extraction from unstructured data sets [78]. This capability and any inference performance associated with DNNs have been widely documented as being highly dependent on the structural configuration of the deep models [77]. Today, the design of more complex neural architectures is expected to show an improvement in model performance, but this design process is both error-prone and time-consuming [**?**, **?**]. Within the last five years, an increasing number of works have focused on studying the application of various methods attempting to automate the process of architectural configuration for DNNs [**?**, 82, 22, 51]. These efforts settled the bases of Neural Architecture Search (NAS), the process of automatically discovering the optimal architecture configuration of a DNN model for a given task.

On the surface, NAS may seem to be easy: to determine the characteristics of an architecture, one needs to build a simple architecture and just let the machine mix and match components until a well-performing model is found. However, NAS involves very large and complex search spaces, and it is implausible to test every, or even most output candidate DNN. Yet, contemporary works on NAS have showcased superior results, outperforming handcrafted DNNs by focusing on maximizing the inference capabilities of synthesized networks [22, 51, 71]. While this situation has allowed the area to consolidate, it has concentrated mostly on object recognition, limiting the discovery of DNN architectures that differ fundamentally from those already studied ad nauseam. We consider NAS is now up to be applied in domains other than image classification.

Image restoration (IR) is a family of inverse problems aiming at obtaining a high-quality image from a corrupted input image [94]. This corruption may occur due to the capture process (e.g., noise, lens blur), post-processing (e.g., JPEG compression), or photography in non-ideal conditions (e.g., haze, motion blur). IR has been well-studied in the area of computer vision and image processing with a relevant importance in many fields ranging from security and surveillance imaging [81], medical imaging [75], to image generation [77]. In the past decades, extensive research has been carried out aiming at developing various IR methods [32]. Nonetheless, deep Convolutional Neural Networks (CNN) have been widely investigated as an approach to IR achieving promising results in various tasks, such as super-resolution image restoration [12, 16, 41, 50, 73], image denoising [4, 15, 31, 42, 76], compressive sensing image recovery [26, 35], and image deblurring [43, 44, 91].

Many deep models that have demonstrated state-of-the-art performance in Super-Resolution Image Restoration (SRIR) tasks have had their results attributed to the networks' ability to learn realistic data priors from large data-sets of images [78]. However, recent studies have discovered that the structure of a network is sufficient in itself to capture rich low-level image statistics. These structured Deep Image Priors (DIP), as the authors have named them in [77], exist subtly encoded in the very structure of a network and are critical for image restoration. This invariably points out that the efficacy of a SRIR model is highly dependent on the architecture used, with different content requiring different architectures to achieve good performance. It is not surprising to find that experts have shown interest in the construction of more precise IR models highly dependable on their architecture and hyperparameters [9, 23, 28].

However, given the lack of a guiding principle to help determine the best performing Deep Learning model architecture to solve a task, discovering an effective architecture becomes a painstaking and slow process that heavily relies on an expert's prior knowledge and experience. Neural Architecture Search (NAS) methods arise as an attempt to automate the process of architecture discovery. Currently, several NAS methods have demonstrated to produce networks that outperform manually-engineered architectures, differing mainly in the search strategies employed by each [19]. Among these strategies, Evolutionary Algorithms (EAs) have gained the attention of experts more recently, due to the computational time required compared to other strategies [19, 21]. These approaches based on evolution have widely become known as neuroevolution, a concept that has resurfaced in the literature in recent years [22, 60, 70].

Contemporary works on neuroevolution have demonstrated outstanding results in the performance of Deep Neural Networks (DNNs) by entirely focusing on maximizing the performing capabilities of the discovered models [22, 71]. However, in real-world applications, the mere consideration of the performance of a model hinders the feasibility of its application, due to limited hardware resources. For example in mobile devices where little available memory and processing power exist capable of running a complex model to solve any IR problem [36]. Another example comes from contemporary works on low-resolution facial recognition where visual data presents heavy degradation. In both cases, not only good performance of the applied models is required, but algorithms still need to improve in their robustness, reliability, and inference time performance [55]

In this and other real-world applications becomes crucial to consider additional objectives while designing a neural model in such cases. The idea of reformulating NAS as a Multi-objective Optimization Problem (MOP) aims at achieving models that require fewer computational resources while maintaining an optimal balance within training speed, accuracy and latency among other aspects. This should allow achieving well-performing yet efficient architectures able to be deployed in situations with resource and time constraints [19, 82]. This approach should also help to study the characteristics of the synthesized architectures contemplating different trade-offs among objectives.

However, few experts have explored the idea of incorporating the concept of Multi-objective Optimisation (MO) while browsing the architecture space. This is motivated greatly by the increase in time complexity and computational resources that the additional objectives bring to the search [22]. The additional computational costs emerge from the fact that no single solution exists for a nontrivial multi-objective optimization problem, but instead a set of Pareto optimal solutions. A solution candidate is considered Pareto optimal if none of the objectives can be improved without degrading some of the other objective values. This consideration changes the final goal of a NAS procedure from synthesizing a single near-optimal model for a task, into that of finding as many models as possible that can be considered Pareto optimal.

Evolutionary Multi-objective Optimization Algorithms (EMOAs) are among the most popular techniques for handling different types of MOPs. Moreover, contemporary works study the incorporation of techniques that help accelerate and even improve an EMOA's approximation to the Pareto front. These techniques aim at improving the searching capabilities of the EMOA, and in some cases, refining its approximation by finding additional solutions missed by the algorithm. This lead to a more frequent adoption of Hybrid EMOAs, broadening the application of these algorithms by combining the global search capabilities of EAs with local search techniques and heuristics, using schemes in which the latter are often subordinated to the former [24].

Our work will explore the design and implementation of a hybrid EMOA for synthesizing deep neural networks, focused on tackling the problem of super-resolution image restoration. This hybrid algorithm should be able to surpass other state-of-the-art NAS approaches in two particular aspects. First, synthesizing efficient SRIR models with better performance than other manually and automatically engineered state-of-the-art models. All this while maintaining a reduction in the computational and time complexity of the models. Second, the incorporation of multi-objective heuristics or local search algorithms aiming at reducing the complexity that the additional objectives introduce. This hybridization should help reduce the overall time and computational cost required for a multi-objective NAS (MoNAS) process to discover models. Especially considering that evaluations involve the training and testing of hundreds and even thousands of Deep Learning models.

1.1 Research questions

The following research questions arise in the context of the information presented above.

- Which measures can be employed to assess the quality of solutions, taking into consideration aspects such as model parametrization, accuracy, memory efficiency and output latency?
- Is it possible to exploit a local search technique to further accelerate the search of effective and efficient architectures?
- Is it possible to further refine the performance and/or efficiency of architectures using a local search technique given a set of near optimal candidates?
- Is it possible to improve the performance of a Multi-objective Neuroevolution approach by hybridizing it using a NAS-focused local search technique?
- In which dimensions (such as performance, efficiency, search latency, etc.) can we expect improvements when comparing our approach to non hybrid multi- and single-objective NAS approaches?

1.2 Structure of the document

WIP...

2 Hypothesis

The hypothesis to be validated in this research is that by the implementation of a hybrid evolutionary approach for multi-objective neural architecture search we will see improvements over different aspects on the discovered final SR architectures. We expect such improvements to be in terms of accuracy, architecture complexity and training speed when contrasted with state-of-the-art manually-engineered and automatically-engineered architectures. Moreover, improvements should also reflect reductions in the computational resources required for the proper deployment of architectures into models, such as memory and processing power.

3 Objectives

The primary objective of this research is to present a multi-objective hybrid neuroevolution approach for MoNAS applied to image SR tasks. This approach will combine an evolutionary algorithm with a local search technique, that should result in a performance improvement obtained by the synthesized networks. Further, the synthesized models should present a reduction in complexity and inference latency.

The more particular goals that will guide this research work are:

1. By the end of 2020, we shall have conducted a systematic review of the state-of-the-art methods in the area of Neural Architecture Search (NAS) with a particular interest in those based on Evolutionary Multi-objective Optimization Algorithms, and so have identified the core elements observable in different approaches.

2. By March 2021, we shall have performed a study of contemporary works on Deep learning approaches to Super-Resolution Image Restoration (SRIR) tasks. Followed closely by an experimental performance contrast among those works that focus on generating models using a MoNAS approach as a way to identify areas of opportunity.

3. A local search algorithm for MoNAS will be developed by early 2022. This algorithm will search in the architecture space for any objective-wise improvement on the neighboring area of the architecture space as well as other Pareto optimal architectures within this same area, allowing the refinement of the solution approximation obtained by a NAS procedure.

4. By the end of 2022, we will have designed a hybrid multi-objective neuroevolution approach that incorporates the previous local search algorithm, for synthesizing near-optimal deep neural models capable of solving super-resolution image restoration tasks.

5. Once the above objectives have been achieved, it is expected that by the end of 2023, an experimental contrast between our algorithm and those of the state-of-the-art will have been carried out. This to identify and highlight the advantages of the algorithm and the resulting models under study.

4 Problem definition and statement

Here we present the concepts, definitions, contemporary areas of interest, and the mathematical models that describe the Neural Architecture Search Problem, as well as, the Super-Resolution Image Restoration problem.

4.1 Neural Architecture Search

A noticeable trend in the field of DL is the construction of more complex DNNs, highly dependable on their architecture and hyperparameters [9, 28, 23]. This has been done to the point that many research efforts in this field focus on finding specialized architectures that excel at a particular task. Nevertheless, given a lack of theoretical understanding of DNNs and no such thing as a guiding principle to help determine the architecture necessary to solve a task, discovering an effective architecture becomes a painstaking process that requires access to massive computational resources and in many cases years of empirical trial and error applications [19, 22, 82].

Neural Architecture Search (NAS) methods arise in an attempt to automate this process of architecture discovery and hyperparameter tuning. Currently, several NAS methods have demonstrated to produce networks that outperform manually-engineered architectures [19]. Different search strategies have been proposed capable of performing NAS, nevertheless, Evolutionary Algorithms (EAs) have gained the attention of experts more recently [19, 21].

Two particular lines of research seem to arise as ways to advance the field of NAS: the application of NAS approaches to new tasks and the use of NAS to reduce the computational cost of neural models in terms of memory usage and processing. Starting with perhaps one of the most commonly criticized aspects of this area is the lack of further exploration of NAS methods in tasks that extend beyond image classification. Some first steps in direction of expanding the domains of application of these methods have been taken in image restoration [30, 72, 83, 94], semantic segmentation [10, 47, 95], and machine translation [68] among others.

Network compression arises in an attempt to reduce the computational costs associated with the deployment of a DL model, especially when resources are limited. In particular, many experts have taken to the task of optimizing complex models for mobile platforms while retaining the same performance as other deeper heavier models. From a NAS approach, this has been explored by limiting the search space, and thus limiting the architectures explored. Few recent approaches explore the possibility of incorporating the limitations in resources and/or time as additional objectives to be optimized either independently or using linear combinations. Even fewer authors have explored the possibility of using Neuroevolution using a multi-objective optimization problem to assert the fitness of models [13, 39, 49, 89].

4.1.1 NAS Problem Statement

Following the works of Wistuba et al. in [82], a general DL algorithm (Λ) can be seen as a function capable of mapping:

$$\Lambda : D \times A \to M \tag{1}$$

where D represents the space of all data-sets, A the architecture search space and M the space of all DL models.

Given a data-set $\delta \in D$, which is split into a training partition δ_{train} and a validation partition δ_{valid}, a general deep learning algorithm Λ estimates the model $m_{\alpha,\theta} \in M$. This is done by minimizing a loss function $\mathcal{L}$ penalized with a regularization term $\mathcal{R}$ with respect to the training data. That is,

$$\Lambda(\alpha, \delta) = \underset{m_{\alpha,\theta} \in M_\alpha}{\arg\min}\ \mathcal{L}(m_{\alpha,\theta}, \delta_{train}) + \mathcal{R}(\theta), \tag{2}$$

where θ is the vector of trainable parameters within the model and $\alpha \in A$ represents the architecture of the model. Following this, a Neural Architecture Search problem is formally defined as:

Definition 1. *(Neural Architecture Search Problem [82])*
Neural architecture search concerns to the task of finding the architecture α^* which maximizes an objective function $\mathcal{O}$ on the validation partition δ_{valid}. In symbols,

$$\alpha^* = \arg\max_{\alpha \in A} \; \mathcal{O}(\Lambda(\alpha, \delta_{train}), \delta_{valid}), \tag{3}$$

where the objective function $\mathcal{O}$ is often defined as the negative loss function $\mathcal{L}$. This is equivalent to solving a black-box global optimization problem for which different approaches exist including the application of EMOAs and Reinforcement Learning (RL) controllers.

4.2 Image Restoration

Image restoration (IR) is an umbrella term that encompasses a family of inverse problems focusing on obtaining a high-quality image from a corrupted input image [94]. IR has been well-studied in the area of computer vision and image processing with a relevant importance in many fields ranging from security and surveillance imaging [81], medical imaging [75], to image generation [77].

In the past decades, extensive research has been carried out aiming at developing various IR methods [32]. Nonetheless, deep CNNs have proved to be an approach to IR capable of achieving promising results in various tasks, such as super-resolution image restoration [12, 16, 41, 47, 73], image denoising [4, 15, 31, 42, 76], compressive sensing image recovery [26, 35], and image deblurring [43, 44, 91].

Recent works, such as that of Ulyanov et al. [77], have shown that CNNs' ability to learn realistic data priors from large data-sets of images is not the only element responsible for the performance of the models, but that the architecture of the models also influences greatly the results. Deep Image Priors (DIP) as the authors named them are low-level image statistics that exists subtly encoded in the architecture of a network and result crucial for the proper recovery and restoration of images.

In many imaging applications, High Resolution (HR) images are often required. HR means that pixel density within an image is high, allowing the image to preserve additional details that could be important for a particular application [27]. Super-Resolution Image Restoration (SRIR) in specific, is a class of image processing techniques in computer vision and image processing that refers to the process of obtaining one or more high-resolution (HR) images from low resolution (LR) images [78].

4.2.1 Super-Resolution IR Problem Statement

This section illustrates the mathematical modelling of a single image SR problem [58, 61]. Such a model represents the degradation process of an observed image, where Low-Resolution (LR) images are acquired from the original High-Resolution (HR) image. A degraded version [17] of the original image is generally obtained due to misfocus of cameras, optical diffraction, comparative motion, storage constraints, improper shuttering, down-sampling, system and environment noise, etc. Then, the main target in SR is to restore a high-quality image from one or several of its degraded versions. Given the particular nature of this process, it is considered an inverse problem with an undetermined inverse equation. One of the commonly used degradation models is presented next, where a LR image is obtained due to blurring, down-sampling, and noise addition.

$$\mathbf{y} = \mathbf{BSx} + v \tag{4}$$

where $\mathbf{B}$ is a blurring matrix, $\mathbf{S}$ is a down-scaling ratio, $\mathbf{x}$ is the original image vector and v is the additive noise vector, resulting in a low-resolution image $\mathbf{y}$. SR methods seek to extract the original image vector from $\mathbf{y}$ by determining the elements that describe $\mathbf{B}, \mathbf{S}$ and v through different means [8, 18, 84]. This represents a very ill-posed and ill-conditioned inverse problem that requires a lot of effort to correctly estimate the original image [6, 7]. Figure 1 represents the different elements that result in a LR image from a HR one and vice versa.

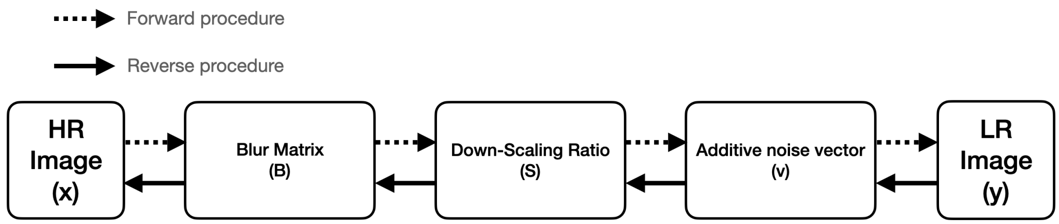

Figure 1: Model for image degradation. The forward procedure shows how a LR image is obtained due to blurring, down-sampling, and noise addition from the original. The reverse procedure requires an algorithm to predict as closely as possible the values of the original image.

5 Related works on the application of NAS to SRIR

Deep learning has provided substantial improvements in the performance of ML within many fields, mostly due to its powerful automatic representation capabilities deeply ingrained in the architecture of models [1, 77, 82]. However, the design of deep neural models heavily relies on researchers' knowledge and experience on a particular task. Interested in the application of Deep Learning techniques with a reduced intervention of experts, and trying to improve even further the performance of models, a myriad of neural architecture search approaches has arisen inside the field of Automatic Machine Learning (AutoML) [19, 22, 82].

Many of these works have managed to achieve competitive results against state-of-the-art manually-engineered models which has demonstrated their effectiveness [22, 64, 71]. Aiming at broadening the application of NAS without the need for specialized hardware, recent trends show that experts are now focusing on discovering simpler resource-efficient networks that retain the same competitive performance of deeper and larger models. This has lead to a recent increment in the works that consider NAS as a multi-objective problem, where aspects like memory size, number of parameters, number of Multiply-Adds operations (Multi-Add), among others, need to be optimized by the models [3, 5, 20, 33, 49, 52, 53, 54, 65, 85, 86, 91, 93]

Not denying the advantages that these approaches bring into consideration, these works have been constantly criticized by the paramount focus that exists on studying the automatic design of object recognition models [20, 52, 54, 85, 86]. This commentary has inevitably pushed NAS research into other areas of interest that include, and are not limited to, object detection and segmentation [48, 74, 96], image inpainting [45], super-resolution image restoration [11, 25], image deblur and denoising [38, 90], network robustness against adversarial attacks [14, 40, 67] and Face Presentation attack detection [87, 88] to mention a few.

The incorporation of more objectives to be optimized and an extensive repertoire of task-dependent architectures, both components resulting in larger and more complex architecture search spaces and models, has driven the experts to consider more sophisticated and complex search strategies [79, 59]. The development of such strategies has deemed many simpler techniques as insufficient for the task of model synthesis and refinement. It was until recently that two works, one by White et al. [79] and one by Ottelander et al. [59], demonstrated that a very simple local search algorithm, hill climbing in both cases, was able to achieve state-of-the-art performance in different popular data-sets.

Both White et al. and Ottelander et al. aim at demonstrating that when the noise found within the evaluation stage of NAS is reduced to a minimum, a trivial local search algorithm is capable of outperforming even popular state-of-the-art NAS approaches [79]. This suggests that many state-of-the-art works have devised complex and sophisticated algorithms to cope with this, rather than attempting to reduce the level of noise in the architecture search space. This demonstrates that there is a propitious line of research for further extend the already promising field of NAS and AutoML [59].

All considerations mentioned above motivated us to follow two lines of research in parallel. First, the application of contemporary EMOAs for synthesizing models applied to super-resolution image restoration, where to the best of our knowledge only a handful of works exist that have considered a multi-objective NAS approach to handle this task [12, 13, 69]. Moreover, these works browse the search space using the NSGA-II, which opens the possibility of studying the applicability and performance of more contemporary EMOAS. Second, the employment of a local search technique used as a continuation method and for the refinement of any discovered architectures. Aiming mainly at improving the velocity in which an EMOA can discover the set of Pareto optimal models and improve the quality of such.

In an attempt at offering a referential frame to highlight the novelty of the contributions that this work seeks, the following sections will present the three contemporary works that have studied the application of neuroevolution MoNAS for SRIR model discovery. These works provide not only a direction to follow for this research but also the possibility of performing a direct comparison of any algorithm and model presented as part of our results.

5.1 Efficient Residual Dense Block Search for Image Super-Resolution

In this paper by Song et al. [69], the authors propose an efficient search algorithm for the discovery of fast, light and accurate networks for image super-resolution. Their models, and thus their search space, are entirely based on residual dense blocks, particularly three efficient cells (blocks) with architectures proposed by the authors, seeking to reduce the computation of networks parameters, like the number of channels, the number of convolution filters, and feature scaling.

Within the evolutionary strategy of their NAS approach, the authors propose the use of information obtained during the training process of each model to measure what they named the credit of a block. The credit of each of the three blocks that conform to the search space is defined as any performance gain by adding a particular block to a model. The credit of a block is then defined as follows:

$$c = f_{add} - f_{bef}, \tag{5}$$

where f_{add} and f_{bef} denote the PSNR of a that includes the new block and before its inclusion, respectively. This block credit is then employed to guide the mutation process of an evolutionary algorithm, where the new architecture of a mutated cell is selected using a roulette, where the probability of selecting a particular dense block is determined by its credits.

To evaluate the quality of solutions, the algorithm employs three particular objectives, the performance of each model measured by the Peak-Signal-to-Noise-Ratio between the resulting images and HR images on the validation set, in conjunction with the number of parameters and number of Multi-Add required for each. This uncommon separation of the number of parameters and Multi-Add comes from the structure of the architecture search space proposed in this work, which considers the inclusion of pooling and recursive operators.

It is crucial to mention that the authors consider two different approaches for their evolutionary strategy. One based on the NSGA-II algorithm, that takes the objectives as part of the same optimization problem and attempts to find the set of Pareto optimal solutions and the other that transforms the problem into a constrained single-objective optimization problem restricting the number of parameters and Multi-Add of a model to always be bellow a particular number. While experimental results demonstrate the effectiveness of the proposed searching method it is not specified whether the authors discovered their resulting objectives by either employing a variation of the NSGA-II and the multi-objective formulation of their problem or a simple genetic algorithm with their constrained single-objective NAS problem.

5.2 Multi-Objective Reinforced Evolution in Mobile Neural Architecture Search

Designing neural models for mobile devices demands additional considerations given the highly constrained resources of such platforms. In this article by Chu et al. [13], the authors introduce a new multi-objective oriented algorithm called MoreMNAS (Multi-Objective Reinforced Evolution in Mobile Neural Architecture Search) combining an EA with RL. In particular, they incorporate a variant of the NSGA-II, where the search space is comprised of many different cells. This means that the crossover and mutation operators used by the EMOA are performed at the cell level.

Moreover, the authors combine the algorithms mutation process with a reinforced control, this way regulating an arbitrary mutation procedure. The idea of combining an EMOA with RL presents an interesting new approach for MoNAS, mitigating the shortcomings of both approaches, as EAs are known to present slow learning capabilities but are more computationally efficient when compared to RL in particular scenarios.

Normally evaluation of models requires the training of any generated model followed by a validation procedure. To compensate for the elevated computational cost of training each of the numerous models found, the authors decided to perform only partial training, based on the empirical conception that better models usually win at early stages. However, less training data allows the introduction of bias.

Here unlike the previous approach, the authors formulate their NAS optimization as a constrained multi-objective problem, still focusing on the maximization of the PSNR for each model while minimizing the number of parameters and number of Multi-Add. The constraints serve to limit the minimum and maximum values that each of the models is supposed to achieve to be considered feasible.

5.3 Fast, Accurate and Lightweight Super-Resolution with Neural Architecture Search

In this work, by Chu et al. [12] the authors extend their previous work, still preserving pretty much the same components. Namely, a cell-based architecture search space, a reinforced driven model generator and a model evaluator based on incomplete training. Differently from the previous work, comes the redefinition of the search space to now combine micro and a macro space on the cell level boosting the model architecture variability during the search process.

Now the search space is designed to allow the NSGA-II algorithm to perform two types of search at the same time. The micro search is used to select promising cells within each block of a model, which can be viewed as feature extraction selection. The macro search then aims at determining the backbone connections for different blocks of the same model, which determines the combination of features extractors selected by the micro search. To avoid increasing the complexity of the encoding and decoding of cells, all blocks share the same search space at the micro-level.

In this work, the authors continue guiding the search by using a constrained multi-objective approach maintaining the same objectives from [13]. However, in this particular case, the constraints only limit the minimum PSNR to improve visual perception of the results and the maximum number of Multi-Add to reduce the resources required for a model to be deployed. Forcing the resulting models to improve in performance while maintaining a reduced number of operations.

6 Methodology

6.1 The Search Space: Branching Architecture Search Space

Neural Networks (NNs) are intricate functions that transform input vectors into outputs through sequential operations. These operations form directed acyclic graphs, each node of which is linked to specific operations such as activation functions, convolutions, and multivariate processes [82, 19]. Neural Architecture Search (NAS) systems are designed to determine the most effective operational combinations, optimizing the sequence and complexity of these graphs to create neural network models tailored for specific tasks [80].

The construction of a NAS's search space is crucial for architectural performance and innovation [19, 56]. This process, encompassing a wide range of design factors, leads to a challenging, high-dimensional optimization landscape. The complexity stems from the extensive array of operations and parameters required to develop a network for specific problems [29, 37]. Constrained search spaces are often favored to manage the architectural synthesis cost, balancing the number of potential candidates while ensuring sufficient variability for effective architecture discovery.

Different approaches, spanning from simple sequential and residual networks to very complex multi-path architectures, offer alternatives for tackling the problem of Super Resolution Image Restoration (SRIR), as discussed in [2]. This work builds on insights presented in prior studies, where authors have shown that multi-branch spaces simplify task learning by reducing duality gaps, enabling simpler architectures with fewer operations to enhance performance. Multi-branch networks, tailored for different aspects of image enhancement like scale and degradation type handling, can capture a wider range of features, significantly improving restoration quality. Their parallel processing capabilities enhance efficiency and speed, while the integration of information from multiple branches results in high-resolution images with greater accuracy and detail. This approach's flexibility to adapt to diverse image data and restoration needs makes it a versatile tool in image processing.

In the following sections, we will delve into the concepts shaping our Branching Architecture Search Space (BASS), illustrating its versatility and potential with examples of resultant architectures.

6.2 Primitive functions: From Basic Operations to Cells in Architecture design

In any Neural Architecture Search (NAS) pipeline, identifying the essential operations that act as building blocks for architecture design tailored to specific problems is critical. These operations, termed *primitives,* are fundamental components that define the range and variability of a search space. By strategically combining these primitives, we can then create *cells* — complex structures that are instrumental in developing advanced neural networks. This process is key to exploring and expanding the potential of neural network architectures within the defined scope of the search space [82, 66].

In this study, we curated a diverse collection of primitives, each with a proven track record in computer vision research. This selection not only aligns with the specific requirements of our task but also introduces operations beyond the traditional scope of super resolution, strengthening the reach of the search space. Our intent is to improve the chances at discovering well-performing architectures while still maintaining enough diversity that may lead to the discovery of novel insights regarding the design of SRIR deep neural networks. The pool of operations for our search space includes a thoughtful mix of these elements.

- **Two-Dimensional Convolution (conv)**: Filters input data using a 2D convolution operation, essential for capturing spatial features in image data.
- **Dilated Convolution (dil_conv_d2, dil_conv_d3, dil_conv_d4)**: Employs dilated kernels to increase the receptive field without additional parameters, used for capturing larger patterns with dilation rates 2, 3, and 4.
- **Depthwise Separable Convolution (Dsep_conv)**: Separates the convolution process into depthwise and pointwise operations, reducing computational load while maintaining effective feature extraction.
- **Inverted Bottleneck Convolution (invert_Bot_Conv_E2)**: Utilizes an expand-transform-project pattern with an expansion rate of 2, optimizing depthwise convolution efficiency for mobile vision networks.
- **Transposed Convolution (conv_transpose)**: Increases the spatial resolution of input features by performing the transposed operation of a conventional convolution, often used in image super-resolution.
- **Identity (identity)**: A no-operation layer that outputs the input as-is, facilitating skip connections and dimension matching in network architectures.

Having delineated the range of primitives available for architectural construction, we now turn our attention to defining a simple feed-forward cell-like structure. This structure is designed to compliment the selection of primitives with variable parameters such as kernel sizes, the frequency of a primitive repetition, allowing a more specific configuration and combination of various primitives as well as increasing the possible lenght of a branch. Additionaly, these parameters include a channel-wide channel configuration such that incompatibility between inputs and outputs is avoided. The integration of these parameters with the selected primitives permits the customization of an architecture at a cell level, catering to specific performance metrics and computational demands. The delimitation of the parametric values that can be searched within the cell structure follows.

- **Channels**: Offering configurations of 16, 32, 48, and 64 to cater to varying levels of feature extraction and width.
- **Repeated Blocks**: Enabling iterations from 1 to 4, which significantly influence the learning capacity and depth of a neural network.
- **Kernel Sizes**: Available in 1, 3, 5, and 7, providing a range of receptive field sizes to accommodate diverse feature detection requirements per layer.

6.3 Branching Architecture Search Space: Head-to-Tail Structured Design

Multi-branch designs seem to be growing in popularity within multiple computer vision tasks with good reason. Zhang et al. in [92] discus how the incorporation of multiple branches helps in reducing the non-convexity of the loss landscape, thereby enhancing the learning efficiency of a network by streamlining the optimization of its weights. Additionally, the advantages of multi-path architectures in super resolution have been examined in several studies, such as [63, 34, 62, 46, 57, 2]. Contrary to sequential or skip-connection based designs, multi-branch networks aim to extract a diverse set of features at various contextual scales. This integration of complementary information results in improved High-Resolution (HR) reconstructions. Furthermore, this design facilitates a multi-path signal flow, enhancing information exchange during the forward and backward steps of training.

Following different examples of multi-branch models for SRIR, our structural framework for BASS begins with a fixed well-defined head cell, with our input passing through a convolutional layer expanding the channels of the original image as needed by the network. Then transitioning into a tripartite backbone, where each branch is composed of different operations, currently limited to only 3 operations per branch selected from the common pool of *primitives* described in

7 subsec: Primitives

where each operation can be repeated multiple times, ensuring distinct yet harmonious functionality. The design concludes with a tail block that aggregates the outputs of the three different branches and proceeds to perform pixel shuffling according to the super resolution requirements of the task to be solved. This tail serves as a critical convergence point, synthesizing the diverse outputs from the branches into a super resolution image. The framework's predefined structural composition not only streamlines operational allocation during the search, but also maintains versatility to adapt the models to various computational budgets and challenges. Figure 2 presents visually the architectural structure of BASS.

Such a branching framework, while featuring a fixed core structure, maintains open the possibility of adaptability and extensibility. As such it allows for the incorporation of an increased number of operations per branch, as well as the potential addition of more branches. Such expansions inevitably enhance the size and scope of the overall search space. This augmentation not only includes a greater variety of combinations using pre-established primitives that can even consider designer blocks, but also facilitates the introduction of additional contextual scales through new branches. Consequently, the BASS can evolve into a more versatile and comprehensive tool for exploring and optimizing multi-path neural network architectures, accommodating a wider array of design possibilities and complexities.

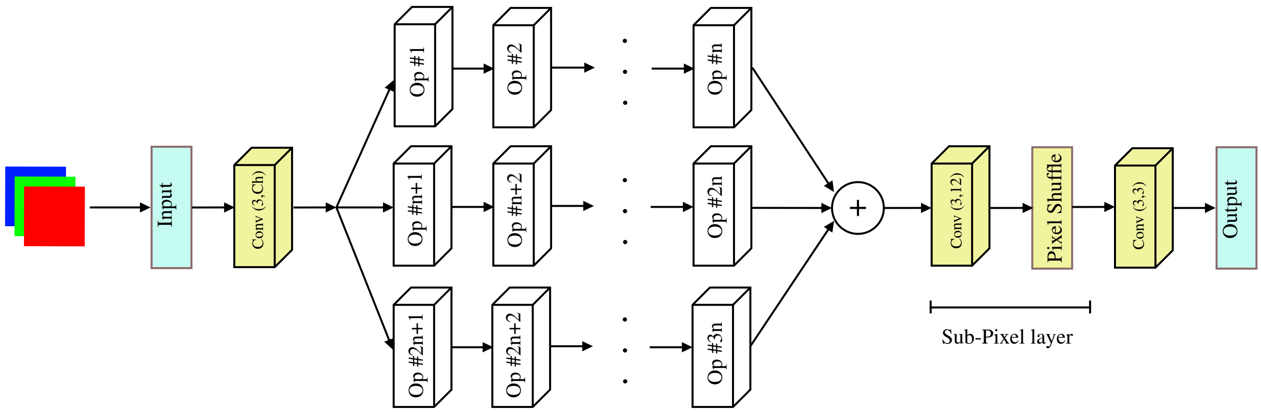

Figure 2: Architecture structure and search space of BASS. The empty boxes represent operations which are selected by the search strategy among the pool of primitives. Conv(3,n) represents a two-dimensional convolution layer layer with n output channels. Operations in yellow blocks are fixed, except for the first Conv in which Ch is determined by a searchable parameter. The n for the purposes of this work is left as 3.

7.1 Going From Artificial Chromosomes to Neural Networks

In our Branching Architecture Search Space (BASS), the way we represent neural networks is crucial. Typically, neural networks are quite complex, being represented as graphs with many connections. This complexity can be challenging to represent withing a Neural Architecture Search (NAS) pipeline. To simplify this, we need a way to represent these networks that makes the search smoother and allows a proper continual browsing of the entire search space and work with. Our encoding strategy in BASS is designed to convert the branching structure found on Section 6.3 and the parameters associated with its searchable backbone into a more straightforward format.

Our approach in BASS uses a special binary encoding method based on Gray codes. This method is beneficial for two main reasons. First, using Gray codes allows us to make small changes in the network architecture that translate in variations between architectures that can be found closer in the search space rather than drastic jumps within different areas. This is because Gray codes are designed to minimize the hamming distance between successive codes. Second, our choice of binary encoding makes BASS more compatible with other existing systems and algorithms. Figure **??** visually presents how a binary string corresponds to a set of parameters and how these are also automatically allocated within the branching structure of our architectural search space.

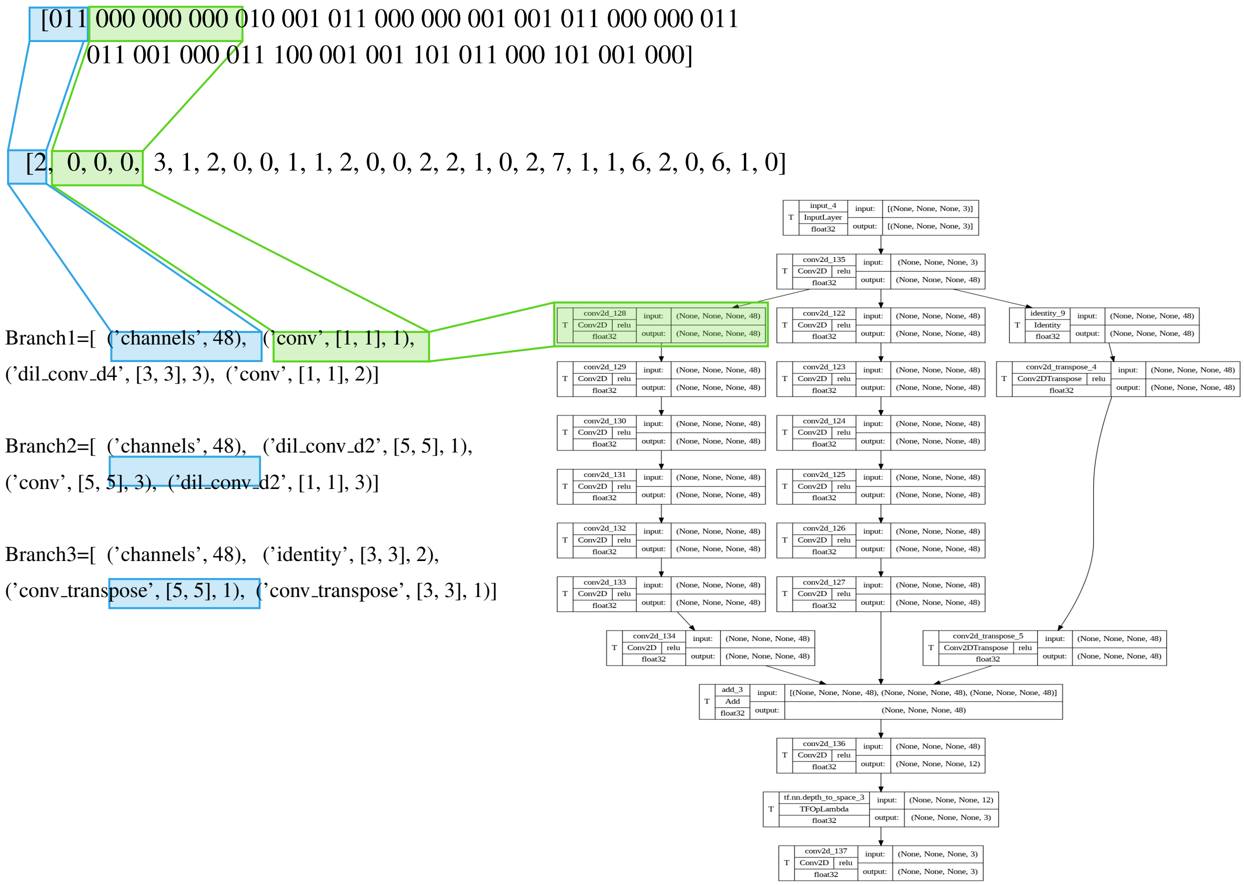

Figure 3: Sequential Mappings: From Binary Gray encoding to Index Assignments, Advancing to Parameter Configurations, and Finally Architectural Structuring in BASS.

References

[1] Md Zahangir Alom, Tarek M. Taha, Chris Yakopcic, Stefan Westberg, Paheding Sidike, Mst Shamima Nasrin, Mahmudul Hasan, Brian C. Van Essen, Abdul A. S. Awwal, and Vijayan K. Asari. A state-of-the-art survey on deep learning theory and architectures. *Electronics*, 8(3), 2019.

[2] Saeed Anwar, Salman Khan, and Nick Barnes. A deep journey into super-resolution: A survey. *ACM Computing Surveys (CSUR)*, 53(3):1–34, 2020.

[3] Maria Baldeon-Calisto and Susana K. Lai-Yuen. AdaResU-net: Multiobjective adaptive convolutional neural network for medical image segmentation. 392:325–340.

[4] Harold C Burger, Christian J Schuler, and Stefan Harmeling. Image denoising: Can plain neural networks compete with bm3d? In *2012 IEEE conference on computer vision and pattern recognition*, pages 2392–2399. IEEE, 2012.

[5] Maria G. Baldeon Calisto and Susana K. Lai-Yuen. Self-adaptive 2d-3d ensemble of fully convolutional networks for medical image segmentation, 2019.

[6] Emmanuel J Candès, Justin Romberg, and Terence Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on information theory*, 52(2):489–509, 2006.

[7] Emmanuel J Candes and Terence Tao. Near-optimal signal recovery from random projections: Universal encoding strategies? *IEEE transactions on information theory*, 52(12):5406–5425, 2006.

[8] Hong Chang, Dit-Yan Yeung, and Yimin Xiong. Super-resolution through neighbor embedding. In *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*, volume 1, pages I–I, 2004.

[9] Zhengping Che, Sanjay Purushotham, Kyunghyun Cho, David Sontag, and Yan Liu. Recurrent neural networks for multivariate time series with missing values. *Scientific Reports*, 8(1):6085, 2018.

[10] Liang-Chieh Chen, Maxwell D. Collins, Yukun Zhu, George Papandreou, Barret Zoph, Florian Schroff, Hartwig Adam, and Jonathon Shlens. Searching for efficient multi-scale architectures for dense image prediction, 2018.

[11] Yun-Chun Chen, Chen Gao, Esther Robb, and Jia-Bin Huang. Nas-dip: Learning deep image prior with neural architecture search. In *European Conference on Computer Vision (ECCV)*, 2020.

[12] Xiangxiang Chu, Bo Zhang, Hailong Ma, Ruijun Xu, and Qingyuan Li. Fast, accurate and lightweight super-resolution with neural architecture search, 2020.

[13] Xiangxiang Chu, Bo Zhang, Ruijun Xu, and Hailong Ma. Multi-objective reinforced evolution in mobile neural architecture search, 2019.

[14] Ekin D. Cubuk, Barret Zoph, Samuel S. Schoenholz, and Quoc V. Le. Intriguing properties of adversarial examples, 2017.

[15] Ratnadeep Dey, Debotosh Bhattacharjee, and Mita Nasipuri. *Image Denoising Using Generative Adversarial Network*, pages 73–90. Springer Singapore, Singapore, 2020.

[16] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Learning a deep convolutional network for image super-resolution. In David Fleet, Tomas Pajdla, Bernt Schiele, and Tinne Tuytelaars, editors, *Computer Vision – ECCV 2014*, pages 184–199, Cham, 2014. Springer International Publishing.

[17] W. Dong, L. Zhang, G. Shi, and X. Li. Nonlocally centralized sparse representation for image restoration. *IEEE Transactions on Image Processing*, 22(4):1620–1630, 2013.

[18] David L Donoho. Compressed sensing. *IEEE Transactions on information theory*, 52(4):1289–1306, 2006.

[19] Thomas Elsken, Jan Hendrik Metzen, and Frank Hutter. Neural architecture search: A survey. *arXiv preprint arXiv:1808.05377*, 2018.

[20] Thomas Elsken, Jan Hendrik Metzen, and Frank Hutter. Efficient multi-objective neural architecture search via lamarckian evolution, 2019.

[21] Dario Floreano, Peter Dürr, and Claudio Mattiussi. Neuroevolution: from architectures to learning. *Evolutionary Intelligence*, 1(1):47–62, 2008.

[22] Edgar Galván and Peter Mooney. Neuroevolution in deep neural networks: Current trends and future challenges, 2020.

[23] Alex Graves, Abdel rahman Mohamed, and Geoffrey Hinton. Speech recognition with deep recurrent neural networks, 2013.

[24] Crina Grosan and Ajith Abraham. *Hybrid Evolutionary Algorithms: Methodologies, Architectures, and Reviews*, pages 1–17. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007.

[25] Yong Guo, Yongsheng Luo, Zhenhao He, Jin Huang, and Jian Chen. Hierarchical neural architecture search for single image super-resolution. *IEEE Signal Processing Letters*, 27:1255–1259, 2020.

[26] Z. Guo, Z. Zhang, and Z. Chen. Deep scalable image compression via hierarchical feature decorrelation. In *2019 Picture Coding Symposium (PCS)*, pages 1–5, 2019.

[27] Viet Khanh Ha, Jinchang Ren, Xinying Xu, Sophia Zhao, Gang Xie, and Valentin Masero Vargas. Deep learning based single image super-resolution: A survey. In Jinchang Ren, Amir Hussain, Jiangbin Zheng, Cheng-Lin Liu, Bin Luo, Huimin Zhao, and Xinbo Zhao, editors, *Advances in Brain Inspired Cognitive Systems*, pages 106–119. Springer International Publishing.

[28] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Identity mappings in deep residual networks, 2016.

[29] Alexandre Heuillet, Ahmad Nasser, Hichem Arioui, and Hedi Tabia. Efficient automation of neural network design: A survey on differentiable neural architecture search. *arXiv preprint arXiv:2304.05405*, 2023.

[30] Kary Ho, Andrew Gilbert, Hailin Jin, and John Collomosse. Neural architecture search for deep image prior, 2020.

[31] Inpyo Hong, Youngbae Hwang, and Daeyoung Kim. Efficient deep learning of image denoising using patch complexity local divide and deep conquer. *Pattern Recognition*, 96:106945, 2019.

[32] Bo Hu, Leida Li, Jinjian Wu, and Jiansheng Qian. Subjective and objective quality assessment for image restoration: A critical survey. *Signal Processing: Image Communication*, 85:115839, 2020.

[33] Junhao Huang, Weize Sun, and Lei Huang. Deep neural networks compression learning based on multiobjective evolutionary algorithms. 378:260–269.

[34] Zheng Hui, Xiumei Wang, and Xinbo Gao. Fast and accurate single image super-resolution via information distillation network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 723–731, 2018.

[35] Nick Johnston, Elad Eban, Ariel Gordon, and Johannes Ballé. Computationally efficient neural image compression, 2019.

[36] Jan Kamenický, Filip Sroubek, Barbara Zitová, Jari Hannuksela, and Markus Turtinen. Image restoration in portable devices: Algorithms and optimization. *Journal of Signal Processing Systems*, 91, 10 2018.

[37] Jeon-Seong Kang, JinKyu Kang, Jung-Jun Kim, Kwang-Woo Jeon, Hyun-Joon Chung, and Byung-Hoon Park. Neural architecture search survey: A computer vision perspective. *Sensors*, 23(3):1713, 2023.

[38] Heewon Kim, Sungyong Baik, Myungsub Choi, Janghoon Choi, and Kyoung Mu Lee. Searching for controllable image restoration networks, 2020.

[39] Ye-Hoon Kim, B. Reddy, Sojung Yun, and Chanwon Seo. Nemo : Neuro-evolution with multi-objective optimization of deep neural network for speed and accuracy. 2017.

[40] Shashank Kotyan and Danilo Vasconcellos Vargas. Evolving robust neural architectures to defend from adversarial attacks, 2020.

[41] Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4681–4690, 2017.

[42] Stamatios Lefkimmiatis. Non-local color image denoising with convolutional neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3587–3596, 2017.

[43] Lerenhan Li, Jinshan Pan, Wei-Sheng Lai, Changxin Gao, Nong Sang, and Ming-Hsuan Yang. Blind image deblurring via deep discriminative priors. *International Journal of Computer Vision*, 127(8):1025–1043, 2019.

[44] Y. Li, M. Tofighi, J. Geng, V. Monga, and Y. C. Eldar. Efficient and interpretable deep blind image deblurring via algorithm unrolling. *IEEE Transactions on Computational Imaging*, 6:666–681, 2020.

[45] Yaoman Li and Irwin King. Architecture search for image inpainting. In Huchuan Lu, Huajin Tang, and Zhanshan Wang, editors, *Advances in Neural Networks – ISNN 2019*, pages 106–115. Springer International Publishing.

[46] Yinhao Li, Yutaro Iwamoto, Lanfen Lin, Rui Xu, Ruofeng Tong, and Yen-Wei Chen. Volumenet: A lightweight parallel network for super-resolution of mr and ct volumetric data. *IEEE Transactions on Image Processing*, 30:4840–4854, 2021.

[47] Chenxi Liu, Liang-Chieh Chen, Florian Schroff, Hartwig Adam, Wei Hua, Alan Yuille, and Li Fei-Fei. Auto-deeplab: Hierarchical neural architecture search for semantic image segmentation, 2019.

[48] Chenxi Liu, Liang-Chieh Chen, Florian Schroff, Hartwig Adam, Wei Hua, Alan Yuille, and Li Fei-Fei. Auto-deeplab: Hierarchical neural architecture search for semantic image segmentation, 2019.

[49] J. Liu, M. Gong, Q. Miao, X. Wang, and H. Li. Structure learning for deep neural networks based on multiobjective optimization. *IEEE Transactions on Neural Networks and Learning Systems*, 29(6):2450–2463, 2018.

[50] Tairan Liu, Kevin de Haan, Yair Rivenson, Zhensong Wei, Xin Zeng, Yibo Zhang, and Aydogan Ozcan. Deep learning-based super-resolution in coherent imaging systems. *Scientific Reports*, 9(1):3926, 2019.

[51] Yuqiao Liu, Yanan Sun, Bing Xue, Mengjie Zhang, Gary G. Yen, and Kay Chen Tan. A survey on evolutionary neural architecture search, 2021.

[52] Mohammad Loni, Sima Sinaei, Ali Zoljodi, Masoud Daneshtalab, and Mikael Sjödin. Deep-Maker: A multi-objective optimization framework for deep neural networks in embedded systems. 73:102989.

[53] Zhichao Lu, Kalyanmoy Deb, Erik Goodman, Wolfgang Banzhaf, and Vishnu Naresh Boddeti. Nsganetv2: Evolutionary multi-objective surrogate-assisted neural architecture search. In Andrea Vedaldi, Horst Bischof, Thomas Brox, and Jan-Michael Frahm, editors, *Computer Vision – ECCV 2020*, pages 35–51, Cham, 2020. Springer International Publishing.

[54] Zhichao Lu, Ian Whalen, Vishnu Boddeti, Yashesh Dhebar, Kalyanmoy Deb, Erik Goodman, and Wolfgang Banzhaf. Nsga-net: Neural architecture search using multi-objective genetic algorithm, 2019.

[55] Luis S. Luevano, Leonardo Chang, Heydi Méndez-Vázquez, Yoanna Martínez-Díaz, and Miguel González-Mendoza. A study on the performance of unconstrained very low resolution face recognition: Analyzing current trends and new research directions. *IEEE Access*, 9:75470–75493, 2021.

[56] Rohit Mohan, Thomas Elsken, Arber Zela, Jan Hendrik Metzen, Benedikt Staffler, Thomas Brox, Abhinav Valada, and Frank Hutter. Neural architecture search for dense prediction tasks in computer vision. *International Journal of Computer Vision*, pages 1–24, 2023.

[57] Abdul Muqeet, Jiwon Hwang, Subin Yang, JungHeum Kang, Yongwoo Kim, and Sung-Ho Bae. Multi-attention based ultra lightweight image super-resolution. In *Computer Vision–ECCV 2020 Workshops: Glasgow, UK, August 23–28, 2020, Proceedings, Part III 16*, pages 103–118. Springer, 2020.

[58] Kamal Nasrollahi and Thomas B. Moeslund. Super-resolution: a comprehensive survey. 25(6):1423–1468.

[59] T. Den Ottelander, A. Dushatskiy, M. Virgolin, and P. A. N. Bosman. Local search is a remarkably strong baseline for neural architecture search, 2020.

[60] Paolo Pagliuca and Stefano Nolfi. Robust optimization through neuroevolution. *PLOS ONE*, 14:e0213193, 03 2019.

[61] Garima Pandey and Umesh Ghanekar. A compendious study of super-resolution techniques by single image. 166:147–160.

[62] Yajun Qiu, Ruxin Wang, Dapeng Tao, and Jun Cheng. Embedded block residual network: A recursive restoration model for single-image super-resolution. In *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 4179–4188, 2019.

[63] Haoyu Ren, Mostafa El-Khamy, and Jungwon Lee. Image super resolution based on fusing multiple convolution neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 54–61, 2017.

[64] Pengzhen Ren, Yun Xiao, Xiaojun Chang, Po-Yao Huang, Zhihui Li, Xiaojiang Chen, and Xin Wang. A comprehensive survey of neural architecture search: Challenges and solutions, 2021.

[65] Christoph Schorn, Thomas Elsken, Sebastian Vogel, Armin Runge, Andre Guntoro, and Gerd Ascheid. Automated design of error-resilient and hardware-efficient deep neural networks.

[66] Simon Schrodi, Danny Stoll, Binxin Ru, Rhea Sukthanker, Thomas Brox, and Frank Hutter. Towards discovering neural architectures from scratch. *arXiv preprint arXiv:2211.01842*, 2022.

[67] Mathieu Sinn, Martin Wistuba, Beat Buesser, Maria-Irina Nicolae, and Minh Tran. Evolutionary search for adversarially robust neural networks. *Safe Machine Learning workshop at ICLR*, 2019.

[68] David R. So, Chen Liang, and Quoc V. Le. The evolved transformer, 2019.

[69] Dehua Song, Chang Xu, Xu Jia, Yiyi Chen, Chunjing Xu, and Yunhe Wang. Efficient residual dense block search for image super-resolution, 2019.

[70] Kenneth O. Stanley, Jeff Clune, Joel Lehman, and Risto Miikkulainen. Designing neural networks through neuroevolution. *Nature Machine Intelligence*, 1(1):24–35, 2019.

[71] Felipe Petroski Such, Vashisht Madhavan, Edoardo Conti, Joel Lehman, Kenneth O. Stanley, and Jeff Clune. Deep neuroevolution: Genetic algorithms are a competitive alternative for training deep neural networks for reinforcement learning, 2017.

[72] Masanori Suganuma, Mete Ozay, and Takayuki Okatani. Exploiting the potential of standard convolutional autoencoders for image restoration by evolutionary search, 2018.

[73] Ying Tai, Jian Yang, and Xiaoming Liu. Image super-resolution via deep recursive residual network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3147–3155, 2017.

[74] Mingxing Tan, Bo Chen, Ruoming Pang, Vijay Vasudevan, Mark Sandler, Andrew Howard, and Quoc V. Le. Mnasnet: Platform-aware neural architecture search for mobile, 2019.

[75] Fuminari Tatsugami, Toru Higaki, Yuko Nakamura, Zhou Yu, Jian Zhou, Yujie Lu, Chikako Fujioka, Toshiro Kitagawa, Yasuki Kihara, Makoto Iida, and Kazuo Awai. Deep learning–based image restoration algorithm for coronary ct angiography. *European Radiology*, 29(10):5322–5329, 2019.

[76] Chunwei Tian, Yong Xu, and Wangmeng Zuo. Image denoising using deep cnn with batch renormalization. *Neural Networks*, 121:461 – 473, 2020.

[77] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. Deep image prior. *International Journal of Computer Vision*, 128(7):1867–1888, Mar 2020.

[78] Zhihao Wang, Jian Chen, and Steven C. H. Hoi. Deep learning for image super-resolution: A survey, 2020.

[79] Colin White, Sam Nolen, and Yash Savani. Exploring the loss landscape in neural architecture search, 2021.

[80] Colin White, Mahmoud Safari, Rhea Sukthanker, Binxin Ru, Thomas Elsken, Arber Zela, Debadeepta Dey, and Frank Hutter. Neural architecture search: Insights from 1000 papers. *arXiv preprint arXiv:2301.08727*, 2023.

[81] W. W. Z. Wilman and P. C. Yuen. Very low resolution face recognition problem. In *2010 Fourth IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, pages 1–6, 2010.

[82] Martin Wistuba, Ambrish Rawat, and Tejaswini Pedapati. A survey on neural architecture search. *arXiv preprint arXiv:1905.01392*, 2019.

[83] Jiangpeng Yan, Shou Chen, Yongbing Zhang, and Xiu Li. Neural architecture search for compressed sensing magnetic resonance image reconstruction. *Computerized Medical Imaging and Graphics*, 85:101784, Oct 2020.

[84] Jianchao Yang, John Wright, Thomas S. Huang, and Yi Ma. Image super-resolution via sparse representation. *IEEE Transactions on Image Processing*, 19(11):2861–2873, 2010.

[85] Zhaohui Yang, Yunhe Wang, Xinghao Chen, Boxin Shi, Chao Xu, Chunjing Xu, Qi Tian, and Chang Xu. CARS: continuous evolution for efficient neural architecture search, 2020.

[86] Weiqin Ying, Kaijie Zheng, Yu Wu, Junhui Li, and Xin Xu. Neural architecture search using multi-objective evolutionary algorithm based on decomposition. In Kangshun Li, Wei Li, Hui Wang, and Yong Liu, editors, *Artificial Intelligence Algorithms and Applications*, pages 143–154. Springer Singapore.

[87] Z. Yu, Y. Qin, X. Xu, C. Zhao, Z. Wang, Z. Lei, and G. Zhao. Auto-fas: Searching lightweight networks for face anti-spoofing. In *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 996–1000, 2020.

[88] Z. Yu, J. Wan, Y. Qin, X. Li, S. Z. Li, and G. Zhao. Nas-fas: Static-dynamic central difference network search for face anti-spoofing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, page 1–1, 2020.

[89] C. Zhang, P. Lim, A. K. Qin, and K. C. Tan. Multiobjective deep belief networks ensemble for remaining useful life estimation in prognostics. *IEEE Transactions on Neural Networks and Learning Systems*, 28(10):2306–2318, 2017.

[90] Haokui Zhang, Ying Li, Hao Chen, and Chunhua Shen. Memory-efficient hierarchical neural architecture search for image denoising, 2020.

[91] Hongguang Zhang, Yuchao Dai, Hongdong Li, and Piotr Koniusz. Deep stacked hierarchical multi-patch network for image deblurring. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.

[92] Hongyang Zhang, Junru Shao, and Ruslan Salakhutdinov. Deep neural networks with multi-branch architectures are less non-convex. *arXiv preprint arXiv:1806.01845*, 2018.

[93] Luna M. Zhang. A new compensatory genetic algorithm-based method for effective compressed multi-function convolutional neural network model selection with multi-objective optimization.

[94] Yulun Zhang, Yapeng Tian, Yu Kong, Bineng Zhong, and Yun Fu. Residual dense network for image restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020.

[95] Barret Zoph, Vijay Vasudevan, Jonathon Shlens, and Quoc V. Le. Learning transferable architectures for scalable image recognition, 2018.

[96] Barret Zoph, Vijay Vasudevan, Jonathon Shlens, and Quoc V. Le. Learning transferable architectures for scalable image recognition, 2018.