Spatially-Adaptive Reconstruction in Computed Tomography using Neural Networks

David Boublil, Michael Elad, Joseph Shtok and Michael Zibulevsky

***Abstract*—We propose a supervised machine learning approach for boosting existing signal and image recovery methods and demonstrate its efficacy on example of image reconstruction in computed tomography. Our technique is based on a local nonlinear fusion of several image estimates, all obtained by applying a chosen reconstruction algorithm with different values of its control parameters. Usually such output images have different bias/variance trade-off. The fusion of the images is performed by feed-forward neural network trained on a set of known examples. Numerical experiments show an improvement in reconstruction quality relatively to existing direct and iterative reconstruction methods.**

***Index Terms*—Computed Tomography, Low-Dose Reconstruction, Neural Networks, Supervised Learning, Filtered-Back-Projection (FBP).**

I. Introduction

Computed tomography (CT) imaging produces an attenuation map of the scanned object, by sequentially irradiating it with X-rays from several directions. The integral attenuation of the X-rays, measured by comparing the radiation intensity entering and leaving the body, forms the raw data for the CT imaging. In practice, these photon count measurements are degraded by stochastic noise, typically modeled as instances of Poisson random variables. There are also other degradation effects due to a number of physical phenomena – see *e.g.* [1] for a detailed account.

Given the projection data, known as the sinogram, a reconstruction process can be performed in order to recover the attenuation map. Various such algorithms exist, ranging from the simple and still very popular Filtered-Back-Projection (FBP) [2], and all the way to the more advanced Bayesian-inspired iterative algorithms (see e.g., [3], [4]) that take the statistical nature of the measurements and the unknown image into account. Since CT relies on X-ray, which is an ionizing radiation known to be dangerous to living tissues, there is a dire and constant need to improve the reconstruction algorithms in an attempt to enable reduction of radiation dose.

In this work we are concerned with the question of image post-processing, following the CT reconstruction, for the purpose of getting better quality CT image, thereby permitting an eventual radiation-dose reduction. The proposed method does not focus on a specific CT reconstruction algorithm, nor the properties of the images it produces. Instead, we take a generic approach which adapts, in an off-line learning process, to any such given algorithm. The only requirement is the access to design parameters of the reconstruction procedure which influence the nature of the output image, such as the resolution-variance trade-off.

We aim to exploit the fact that any reconstruction algorithm can provide more image information if instead of one fixed value of a parameter (or a vector of them) controlling the reconstruction, few different values are used (leading to different versions of the image). In order to extract this information from a collection of image versions, we use an Artificial Neural Network (ANN) [5]. The proposed method can also use other techniques for computing a non-linear multivariate regression function.

Neural networks have been used extensively in medical imaging, particularly for the purpose of CT reconstruction (see Section III for an overview). Here we propose a new constellation, which consists in a local fusion of the different image versions, aimed at an improved reconstruction quality. We use a set of intensity values from a neighborhood of a pixel q, taken from the different versions, as inputs to the network, and train it to compute a (smaller) neighborhood of q which values are as close as possible (in Mean-Squared-Error or other sense) to those found in the reference image. As we show in this paper, the proposed approach enables an improvement of the variance-resolution trade-off of a given reconstruction algorithm, i.e. producing images with a reduced amount of noise without compromising the spatial resolution and without introducing artifacts.

This paper is organized as follows: Sections II and III are devoted to a brief and general discussion on CT scan/reconstruction and artificial neural networks. Readers familiar with these topics may skip and start reading at Section IV, where the core concept of this work is detailed. In the sequel, the proposed method is implemented on two tomographic reconstruction methods: boosting the Filtered Back-Projection (FBP) is presented in Section V and the same for Penalized Weighted Least-Squares (PWLS) method is described in Section VII. We conclude this work by discussing the computational complexity of the proposed algorithms in Section VIII, and a summary of this work and its potential implications in Section IX.

II. Background on Computed Tomography

In this section we briefly describe the CT imaging and reconstruction processes, assuming two dimensional tomography, and set notations for the rest of the paper.

All authors are with the Computer Science Department, Technion - Israel Institute of Technology, Israel.

The research leading to these results has received funding from the European Research Council under European Union's Seventh Framework Program, ERC Grant agreement no. 320649. This research was supported by Gurwin Family Fund.

From the mathematical point of view, in CT imaging the object is a function $f(x)$ in the plane, which values are the attenuation coefficients of composing materials. When the measured photon counts are ideal, the measurements are directly related to the X-ray transform of $f(x)$ as a collection of all the straight lines passing through the object, and the value associated with each such line is the integral of $f(x)$ along it. Under the assumption of a full rotated and parallel beam scan, this coincides with the Radon transform $\mathbf{R}f$.

Let ℓ be a straight line from an X-ray source to a detector. The ideal photon count λ_ℓ, measured by the detector is related to $\mathbf{R}f$ via the function

$$\lambda_\ell = \lambda_0 e^{-[\mathbf{R}f]_\ell}, \tag{II.1}$$

where λ_0 is the blank scan count. The scanned data is stored in a matrix which columns (referred to as projections) correspond to the sampled angle θ; and are acquired, by a parallel array of X-rays passing through the object at the corresponding angle. According to (II.1), for reconstruction purposes the measurements data undergoes the log transform

$$g_\ell = -log\left(\frac{\lambda_\ell}{\lambda_0}\right). \tag{II.2}$$

Since g, referred to as the sinogram matrix, is the sampled Radon transform of the original image, a discrete version of it can be reconstructed by applying the inverse Radon transform.

Each measured photon count y_ℓ is typically interpreted as an instance of the random variable Y_ℓ following a Poisson distribution $Y_\ell \sim Poisson(\lambda_\ell)$ [1], [6], [7].This reflects the photon count statistics at the detectors [8].

There are various reconstruction algorithms that aim at computing the attenuation map of the scanned object from its projections. In this paper we shall refer and work with two such algorithms: (i) the Filtered Back-Projection (FBP) [2], which is a direct Radon inversion approach; and (ii) an iterative Bayesian reconstruction algorithm that takes the statistical nature of the unknown and the noise into account (e.g. [3]). Bayesian methods achieve better image quality than the direct Radon inversion, at the expense of longer processing time.

Filtered-Back-Projection: This is a linear operator of the form

$$\mathbf{T}_{\text{FBP}} = \mathbf{R}^*\mathbf{F}_{low}\mathbf{F}_{RL}. \tag{II.3}$$

$\mathbf{R}^*$ is the adjoint of the Radon transform (back-projection). $\mathbf{F}_{RL}$ is a 1-D convolution filter, applied to each individual projection, using the Ram-Lak kernel [9], defined in the Fourier domain by $\hat{k}(\omega) = |\omega|$. $\mathbf{F}_{low}$ is a low-pass filter which prevents noise amplification at high frequencies, typical for the Ram-Lak action. Without low-pass filter, the FBP is an exact inverse of the Radon transform in the continuous domain [10] for the noiseless case. Moving from theory to practice, the FBP algorithm does not perform very well. The low-pass filter in the sinogram domain is not an effective remedy for the projections noise. The problem of photon starvation manifests through outlier values in the sinogram, which propagate to the output image in the form of streak artifacts. They corrupt the image contents and jeopardize its diagnostic value.

Statistically-Based Approach: the relation between f, the desired CT image, and the vector of measured counts y can be described as

$$\log\left(\frac{y}{y_0}\right) = \mathbf{A}f + e, \tag{II.4}$$

where $\mathbf{A}$ approximates the Radon transform and models the scan process in reality, and y_0 is the measured photon counts without any attenuation. The additive error e (which depends on f) stems from the statistical noise. In the Bayesian framework, the reconstruction is performed by computing the Maximum a-Posteriory (MAP) estimate of the image

$$\tilde{f} = \arg\max_f \mathbf{P}(f|y) = \arg\max_f \frac{\mathbf{P}(y|f)\mathbf{P}(f)}{\mathbf{P}(y)}. \tag{II.5}$$

For CT, an accurate statistical model of the data is complicated and is often replaced by a Gaussian approximation with a suitable diagonal weighting term whose components are inversely proportional to the measurement variances. This leads to a penalized weighted least-squares (PWLS) formulation [11]:

$$\tilde{f} = \arg\min_f \left\|\log\left(\frac{y}{y_0}\right) - \mathbf{A}f\right\|_D + \beta R(f), \tag{II.6}$$

where $\|u\|_D = u^T\mathbf{D}u$, $\mathbf{D}$ is a diagonal matrix of weights, which in simplistic model are proportional to photon counts y. The penalty term $R(f)$ also referred to as the *prior*, expresses assumptions on the behavior of the clean CT image. In [12] this expression is chosen as

$$R(f) = \sum_q \sum_{k\in\mathcal{N}(q)} \psi_\delta(f_q - f_k), \tag{II.7}$$

where for each location q, a scalar function $\psi_\delta(x)$ is the convex edge-preserving Huber penalty

$$\psi_\delta(x) = \left\{\begin{array}{cc} \dfrac{x^2}{2}, & |x| < \delta \\ \delta|x| - \dfrac{\delta^2}{2}, & |x| \geq \delta \end{array}\right\},$$

L-BFGS optimization method [13] is used to minimize (II.6).[1]

III. Artificial Neural Networks (ANN)

For completeness of this paper, we provide here a brief background on ANN, and in particular their role in CT and medical imaging. ANN, mimicking after the biological networks of neurons which comprise the nervous system, are intensively used in many domains of Computer Science. In this work we focus on the multi-layer feed-forward ANN with no cycles. This is best represented by a directed, weighted graph which has an array of input nodes (data inputs), inner nodes (neurons) implementing specific (linear or non-linear) scalar functions, and another array of output nodes. Input argument of each neuron is the weighted sum of all its inputs, where the weights are associated with the edges. Those weights are learned during the network training and, effectively, define the regression function produced by the ANN.

[1] A Matlab/C code implementation of this algorithm is found in ...., courtesy of Mark Schmidt.

More specifically, the first layer consists of m inputs, coming from the outside world; then N_l neurons are situated in the l-th layer ($l > 1$), and the last one contains n output nodes. Each input x_i is connected to each neuron j in the second (hidden) layer by a weighted edge with weight $w^1_{i,j}$. The output of each neuron is connected to the input of every k-th neuron in the second layer by the weight $w^2_{j,k}$, and so on. Finally, each neuron of the last layer is connected to the output y_s with a weight $v_{s,j}$. We denote by σ the function implemented in each neuron. There is a number of popular choices for this function, for instance $\sigma(x) = \tanh(x)$.

For example, here is the explicit definition of a network with one hidden layer:

$$y(x; w, v, b) = \sum_j v_j \sigma \left(\sum_i w_{i,j} x_j + b_j \right). \tag{III.1}$$

The weights $\{w, v, b\}$ define the multi-variable regression function $y = y(x)$ which approximates any continuous function implied by the set of training examples[2]. A training set for the network comprises of a collection of examples (X^k, Y^k), where X^k is the vector of inputs and Y^k is the true output related to this vector. Training the network consists of optimizing the weights $\{w, v, b\}$ for a minimal error,

$$(w, v, b) = \arg\min_{w,v,b} \sum_{k=1\ldots K} E\left(y(X^k; w, v, b), Y^k\right), \tag{III.2}$$

where the sum is over the training set, and E(a,b) is an error measure of some sort (e.g. $E(a, b) = (a - b)^2$). The popular method for solution of this problem is the iterative backpropagation method [15]. A scheme of such network is depicted in Figure 1.

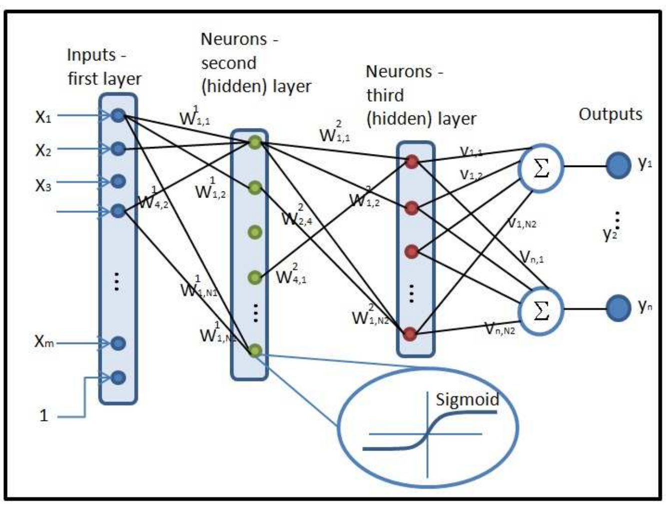

Figure 1: A scheme of a multi-layer feed-forward ANN.

Since the development of the back-propagation algorithm for ANN in mid-eighties, the image processing community (among others) has attained a powerful tool to attack virtually any regression or discrimination task. Among the wealth of applications neural networks found in this area (see [16] for a broad and comprehensive overview), some were designed for medical imaging. As such, Hopfield ANN were used for computer-aided screening for cervical cancer [17], breast tumors [18] and segmentation [19]. ANN are also used for compression and classification in cardiac studies [20] and ECG beat recognition [21]. Tasks of filtering, segmentation and edge detection in medical images are addressed with cellular ANN in [22]. Our group has used neural networks for optimal photon detection in scintillation crystals in PET [23].

As for reconstruction problems, a series of works has appeared in which the ANN replaces the overall reconstruction chain by learning the net contribution of all detector readings to each pixel in the image. For Electron Magnetic Resonance (EMR), such an algorithm is proposed in [24]. Floyd et. al. have used this approach for SPECT reconstruction [25] with feed-forward networks and also for lesion detection in this modality [26]. We remark that such naive application of the ANN for reconstruction is limited to low-resolution $n \times n$ images, since the network must have $\mathcal{O}(n^2)$ inputs and outputs. For instance, in [24], a 64×64 image is reconstructed. Application of ANN for SPECT reconstruction was also studied by J. P. Kerr and E. B. Bartlett [27], [28].

Imaging modalities like PET and SPECT, where low-resolution images are produced, are a natural domain for ANN application. However, some works tackle also the problem of CT reconstruction where the image size is larger. Ref. [29] proposes using a neural network structure with training based on a minimization of a maximum entropy energy function. Reconstruction in Electrical Impedance Tomography was treated with ANN in [30]. Another variety, an Electrical Capacitance Tomography and an ANN-based reconstruction method for it, are described in [31].

Despite the abundance of applications, there is still place for innovation in the domain of ANN application for medical imaging. First, the CT reconstruction problem is rarely approached with this tool due to the high dimensions of raw data and the resulting images, which render the naive application of ANN as the black box converting measurements to image unfeasible. Indeed, in our work we do not propose such a scheme per se – rather, our ANN is employed to perform a locally-adaptive fusion of a number of image versions, produced by a given reconstruction algorithm upon using different configurations. This brings us naturally to the next section where we describe our algorithm.

IV. The Proposed Scheme

A. Local Fusion with a Regression Function

We consider the general setup of the non-linear inverse problem. Assume we are given the measurements vector y of the form

$$y = \mathbf{H}x + \xi, \tag{IV.1}$$

where $\mathbf{H}$ is some transformation, ξ represents the noise, and x is the signal to be recovered. Assume further that $\mathbf{T_p}$ is some

[2] The Universal Approximation Theorem states that a network with just one hidden layer, where each neuron is realized as a monotonically-increasing continuous function, can uniformly approximate any given multivariate continuous function up to an arbitrary small error bound [14]. In practice, adding hidden layers shows an improvement in the ANN performance.

restoration algorithm designed to recover x from this type of measurements, i.e.,

$$\bar{x}_{\mathbf{p}} = \mathbf{T}_{\mathbf{p}}(y) \tag{IV.2}$$

The scalar parameter $\mathbf{p}$ controls the behavior of $\mathbf{T}$ and therefore influences certain characteristics of the estimate $\bar{x}$. For example, when $\mathbf{p}$ is responsible for variance-resolution tradeoff of the algorithm, the estimate $\bar{x}_{\mathbf{p}}$ may be obtained with different noise levels and corresponding spatial resolution characteristics.

The described situation is common in many signal/image processing scenarios. As a basic example, we consider a simple image denoising algorithm, which recovers the signal x from noisy measurements $y = x + \xi$ by a shift-invariant low-pass filter, realized as a 2-D convolution with prescribed kernel. For some fixed shape of this kernel (say, a simple boxcar function or a 2-D Gaussian rotation-invariant kernel), its width (spread) can be parameterized by a scalar variable $\mathbf{p}$. A wider such kernel will perform a more aggressive noise reduction, by averaging the noisy signal over a larger area, at the cost of reducing the spatial resolution.

A second, and more relevant example to this work, is from the domain of CT reconstruction. Recovery of the attenuation map is classically performed by the Filtered Back-Projection algorithm. The latter involves a 1-D low-pass filter, applied to the individual projections. As in the above example, the cut-off frequency of this filter controls the variance-resolution properties of the reconstructed image. In these examples, and also in a general such situation, no single value for the parameter $\mathbf{p}$ makes the best of the processing algorithm. For different signals, different values may be optimal in the sense of MSE or other quality measure. Indeed, in the same image, computed with two different values of $\mathbf{p}$, different regions will get the best treatment by different values of $\mathbf{p}$. For each specific case, ad-hoc considerations for tuning this scalar parameter are applied.

In the domain of non-parametric statistics, there is a noise reduction algorithm with proven near-optimality that devises a switch rule for selecting at each location of the signal an appropriate local filter [32]. In effect, the signal is processed by a low-pass filter adaptive to the local signal smoothness. In the context of our discussion, one can say that this algorithm performs a fusion of a number of filtered versions of a signal with varying filter parameter. The switch rule, developed for this adaptive signal smoothing, is based on the balance of the stochastic and structural noise components and model assumptions, and as such, it is very difficult to devise. Moreover, better output may be obtained if we allow to use some combination of the given image versions in each pixel, rather than selecting one of them alone. To our knowledge, no mathematical theory offering a descriptive rule for such local fusion is available for signal estimators, used for denoising or CT reconstruction.

Borrowing from the above switch-rule idea between filters, the solution we propose for the problem described above is a local fusion of a sequence of estimates $\bar{x}_{\mathbf{p}_1}, ..., \bar{x}_{\mathbf{p}_N}$ with a specific regression function, learned on a training dataset consisting of similar cases. Among known regression methods, we choose to work with ANN, due to their strong adaptivity and generalization properties [5]. The supervised learning is done with a training set of examples: For each location in the processed signals, the features (input vector) are sample values extracted from the corresponding location in the sequence of reconstructed versions for this signal. The output is a small region of sample in the desired destination signal. Contemporary training algorithms employ error back-propagation to optimize the objective function, measuring the discrepancy between the correct output values and those predicted by the ANN [15]. In our work we employ the Caffe Software [33]; the training is performed using the Stochastic Gradient Algorithm. We have tested several sizes and depths of ANN and report here only some of our experiments. We use the function $\sigma(x) = tan(x))$ as the activation function.

In this work, the outlined general concept is specialized to reconstruction algorithms for CT. Specifically, we consider representatives of the two types of those algorithms: the direct FBP and the iterative PWLS (Section II) methods. For FBP, we propose making a sweep over the cut-off frequency of its low-pass filter in the sinogram domain. This parameter controls the noise-resolution tradeoff and has a major influence on the visual impression of the resulting images. For the iterative PWLS algorithm, a sequence of images is extracted along its execution by saving a version of the CT result every few iterations. Along the way, we discuss the choice of training set and design of features extracted for the ANN.

B. *Error Measures*

Just before we conclude this section and move to present the specific details of boosting CT reconstruction algorithms, we should discuss the choice of the error function to use in the learning process, and the error measure to use when evaluating the quality of the reconstruction.

Training Error

In the supervised learning procedure, we design the ANN weights so as to minimize the regression error between the ANN output and the desired outputs (training output data). In many cases, the natural choice for this function would be the Mean-Squared-Error (MSE). However, in CT, we should contemplate whether MSE is the proper choice to use. Consider a homogeneous region in a CT image (corresponding to some tissue) with a small detail of a different yet similar intensity (a cavity or a lesion). The MSE penalty paid by an over-smoothing reconstruction filter that blurs this lesion is small, and therefore such faint details may be lost while leading to better MSE. The remedy for this problem could be to penalize not only for the difference in intensity values between the reference image f_0 and the reconstruction $\hat{f}$, but also for the difference in the derivatives of these two images. Alternatively, we can weight the training examples so as to boost the importance of such faint edge regions, at the expense of more pronounced parts of the image, where the edges are sufficiently strong. This is a simple weighted MSE. In our work we chose to work directly with the MSE objective despite the above delicate matter, leaving the treatment of faint edges as future work direction. We should note that we do take into account the existence of many "empty" patches that are uninformative and may divert the training process. We handle this by randomly choosing a certain amount of examples among a very large set, with probability proportional to the accumulated gradient value. This way, the training is made on the most informative patches.

Quality Assessment

The quality measures of CT images used in this study, are the following:

- **Signal-to-Noise Ratio** (SNR), defined for the ideal signal f and a deteriorated version $\hat{f}$ by $\text{SNR}(f,\hat{f}) = -20log_{10}(\|f-\hat{f}\|_2/\|f\|_2)$. In practice, we consider the signal $\hat{f}$ up to a multiplicative constant and compute
$$SNR(f,\hat{f}) = \max_{\alpha} -20log_{10}(\|f-\alpha\hat{f}\|_2/\|f\|_2). \quad \text{(IV.3)}$$
To make the error measurement more meaningful, the SNR is only computed in the image region where the screened object resides, ignoring the background area. We have used an active contour technique to find the object region in the image; specifically we have used the Chan-Vese method [34].
- **Windowed Signal-to Noise Ratio**. The dynamic range of the HU values in a CT image is very large, from -1000 for air to $1500-2000$ for bones. Often, the diagnostic interest lies in the soft tissues, the HU values of which are near zero (HU of water). For axial sections of thighs, we chose (by a criterion of best visibility) the window of $[b_1=-220, b_2=350]$ HU; our algorithms are tuned for best reconstruction in this HU range. Therefore, an appropriate SNR measurement considers only the regions in the image that fall in this range. Technically, the reference image f and the noisy image $\hat{f}$ are pre-processed before the standard SNR is computed by projecting values lower or higher than b_1 and b_2 respectively to these values.
- **Structured Similarity** (SSIM) measure [35]. This measure of similarity between two images comes to replace the standard Mean Squared Error (the expression $\|f-\hat{f}\|_2$ appearing in the SNR formula), which is known for its problamatic correlation with the human visual perception system (see [35] and the references 1-9 therein). SSIM compares small corresponding patches in the two images, after a normalization of the intensity and contrast. The explicit formula involves first and second moments of the local image statistics and the correlation between the two compared images. In our numerical experiments, we use the Matlab code provided by the authors of [35], which is available at https://ece.uwaterloo.ca/∼z70wang /research/ssim.

V. FBP Boost – Algorithm Design

A. The Scanning Process Model

The ANN needs to be trained on examples which will define the regression inputs and targets. These examples should be composed of several pairs of high-quality images and corresponding corrupted sinograms. In order to build these pairs, original clean CT images are taken (corresponding to the attenuation map of scanned bodies), and the scanning process model is simulated on them to produce corrupted sinograms. The reconstruction algorithms described above are used to get the corresponding ”corrupted images”. The sinogram is obtained by applying a discrete Radon transform on the original image, the resulting matrix is used to compute the expected photon count according to Equation II.1, this way we obtain the ideal photon count λ_ℓ. Then each photon count y_ℓ is generated using the Poisson distribution $Y_\ell \sim Poisson(\lambda_\ell)$ and finally a white Gaussian noise is added. By controlling the levels of the noise, we tune the level of corruption in our raw data.

B. The Low-Pass FBP Filter Parameters

The method of local fusion, advocated in the previous section, is now applied to the standard Filtered Back-Projection (FBP) algorithm for CT reconstruction. The fusion is performed over the parameters of the low-pass sinogram filter, applied before the Back-Projection. This one-dimensional low-pass filter is realized as a multiplication with the Butterworth window H in the Fourier domain, defined by

$$|\hat{H}(\omega)| = \left(1+\left(\frac{\omega}{\phi_0}\right)^{2p}\right)^{-1/2}. \quad \text{(V.1)}$$

We sweep through the range of the parameter ϕ_0 (expressing the cut-off frequency of the filter), thus changing the resolution-variance tradeoff of the FBP. We also change the parameter p, which controls the steepness of the window roll-off. While ϕ_0 controls the amount of blur introduced during the reconstruction, the parameter p influences the texture of reconstructed image.

In Figure 2 we show the reconstruction for a fixed value of $p=3$ and an increasing cut-off frequency ϕ_0. Visually, the strong low-pass filter produces a cleaner image (which also have a higher SNR), but looses in the spatial resolution. The displayed sequence corresponds to values $\phi_0 = [0.4, 0.8, 1.15, 2.0, 120, \infty]$ (the last corresponds to no filter).

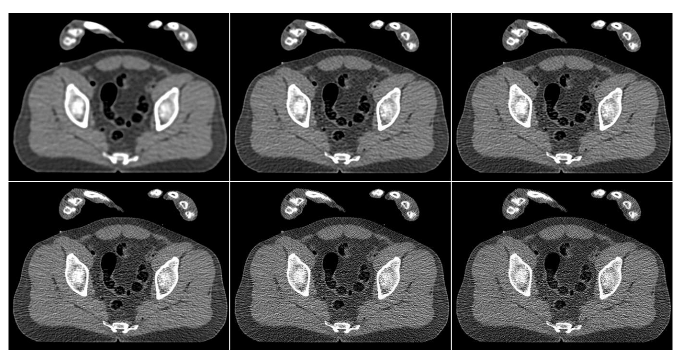

Figure 2: FBP reconstruction with different cut-off frequency value. Top to bottom, Left to right: $\phi_0 = [0.4, 0.8, 1.15, 2.0, 120, \infty]$ (the last image is compute without the low-pass filter).

In order to perform an effective fusion we have to choose several FBP images with different characteristics, some blurier and some noisier but sharper, this way the ANN will smartly merge these and benefit from the best properties of all of them. In this context, the ANN can be regarded as a powerful non-linear/adaptive filter applied to all the FBP images, such that the local fusion is dependent on the specific patch being processed. In order to define coarsely the filters cut-off frequencies, the projections frequency content is analysed on several images, however the fine-tuning is performed empirically: visually and quantitatively. After testing various combinations, we chose to use only three FBP images with cut-off frequencies $\phi_0 = [0.4, 2, \infty]$ and $p = 3$. The first one will be the blurier image as the filter is more agressive; the last is not using any filter and thus it is expected to be noisier but sharper. The middle option gives a good compromise (and will generally get better performances) between the two. Those were selected from eight images – three with the frequencies $\phi_0 = [0.4, 0.8, 1.15]$ and $p = 1$, other three images with the same frequencies and $p = 3$, and the last two are obtained with $\phi_0 = [2.0, 120]$ and $p = 3$. The reason for the restriction to three images arises from the ANN size restriction.

C. Design of The ANN Fusion and Training Setup

Let $\tilde{f}_1, ..., \tilde{f}_K$ be a given set of versions of a CT image, reconstructed by FBP with different low-pass filters in the sinogram domain.[3] We describe the fusion procedure used to compute the output image $\hat{f}$ of the algorithm:

- For each location q in the image matrix, extract its patch [4] neighborhood from each of the K images $\tilde{f}_i$, $i = 1, ..., K$. For our experiments the patch size is set to 7×7.
- Compose a set of inputs for the ANN by stacking the pixel intensities from the K neighborhoods into one vector. Normalize this vector in the training stage (discussed below).
- Apply the ANN to produce a set of output values, which are the intensity values in the patch neighborhood of q in the image $\hat{f}$. The output patch size is set to 3×3.
- By this design, each pixel in the output image is covered by 9 patches; its final value is computed by averaging all those contributions.

We detail now on the several of the steps in the list above. In the training stage, the neural network is tuned to minimize the discrepancy between true values in each output vector and those produced by the network from the set of noisy inputs. A vector of inputs is built, as described above, for a location q in a reference image f from a training set, using data from noisy reconstructions. The corresponding vector of outputs is the patch neighborhood of q in the reference image. Thus, for each image f we produce the set $\tilde{f}_1, ..., \tilde{f}_K$ (as said above, k=3 in our case) using pre-defined FBP filters and sample them to build the training dataset. The image is sampled on a cartesian grid, choosing every second pixel q both in horizontal and vertical directions. The pair of input and output vectors for the neural networks is an example used in the training process. Examples from all the training images f are put in one pool. Among all these examples $30,000$ are randomly chosen, a higher probability of selection is given to patches containing higher gradients on average. This way the too smooth (and non useful) patches are discarded with high probability, as these generally correspond to regions of air (since no constant patch in any kind of tissue can be observed in the noisy FBP images), and the training is mainly performed on high variance patches. This step leads to an empirical improvement in the performance of the ANN.

It is generally acknowledged, that data normalization improves performance of neural networks [36]. Our data matrix A, which columns are the individual example vectors, is normalized by

$$A \Leftarrow A - \min_i(A(i)) \text{ and then } A \Leftarrow A / \max_i(A(i)). \quad \text{(V.2)}$$

The two constants $\alpha_1 = \min_i(A(i))$ (the minimum value of the matrix A) and $\alpha_2 = 1/\max_i(A(i))$ are stored along with the weights of the neural network, and the new data matrix in the test stage is transformed with those pre-computed constants.

Given intensity values in the neighborhood of a pixel q in several noisy images, the network should predict a single value in this pixel for the fusion image. However, as a step of regularization, we design the ANN to produce a vector output which is interpreted as a small neighborhood of q. the fusion image is then built from such overlapping patches, which are averaged to produce the final result. This is done to avoid possible artifacts, which can be produced by the network: in the training stage, if the ANN produces a single outlier intensity value, its penalty will be smaller than of a vector of such incorrect intensities. Such regularization reduces the performance the ANN can achieve on the training set, since more equations are imposed, but its performance on test images is expected to be more stable.

A summary of the whole process: training data preparation, ANN training and application is given in Algorithm 1.

VI. FBP Boost – Empirical Study

In this section we show emperical results of the FBP boosting. In the first part we present some results obtained when the fusion is applied on several images, then we show that even when applying the ANN on a single FBP version (single image fusion), a significant improvement is obtained.

A. Evaluating the Algorithm Performance

In the experiments we have used sets of clinical CT images, axial body slices extracted from a 3D CT scan of a male and female head, thorax and abdomen. The images are courtesy of the Visible Human Project [37]. The intensity levels of those grayscale images correspond to Hounsfield Units. The training set comprises of twelve 461×461 male abdomen sections. The test is performed on four different images taken from the male abdomen (about 10 cm away from the region from which the training data was taken), thorax and head and from the female abdomen. We have performed experiments using different sizes and depths of ANN, and we shall focus our report on two specific examples: a 3 hidden layers ANN and a 7 hidden layers ANN, each comprises of 20 neurons.

[3] Note that all these images are produced from the very same raw sinogram, which means that the patient is exposed to radiation only once.

[4] Note that we use square patches, but disk-shaped patches may be used as well.

Algorithm 1 ANN boost of FBP.

1) Create training data based on M clean CT images:
   For every training image:
   a) Simulate low count CT sinogram with combined Poisson and Gaussian noise.
   b) Reconstruct the image from the simulated sinogram using FBP with K low-pass filters.
   c) From the set of all overlapped patches randomly extract a subset of patches for training, with probability proportional to the average gradient norm in the patch (in order to better treat edges).
   d) Central regions of the corresponding clean patches are extracted to form output training data.
   e) Normalize the data according to equation V.2.
2) Train ANN with e.g. stochastic gradient descent using the above training data.
3) Image fusion with ANN:
   a) Perform K reconstructions with the same low-pass filters as in the training.
   b) ANN application: for each pixel do
      - Normalize the data according to equation V.2.
      - Apply trained ANN.
      - De-normalize using stored α_1 and α_2 (V.2).
      - Position each patch in its place: each estimated pixel is computed by averaging over all the contributions.

For the shallower ANN, $30,000$ patches are extracted to be used as training set, while for the deeper one $100,000$ patches are used. These quantities suffice to avoid over-fitting for the chosen sizes of neural networks. The vector of features for each example is built from the pixel neighborhood of size 7×7, coming from the three corresponding FBP reconstructions. We have used this input patch size as it is leading to the best results while using a reasonable amount of training examples and iterations. As said before, these input images are a subset of the 8 FBP reconstruction results, seeking for the subgroup that would perform the best. The size of the input vector is thus $3 \times 49 = 147$ entries. The output vector is built from the pixel neighborhood of size 3×3 of the clean version.

In Figures 3 and 4 we present a reconstruction of 2 different test images: female abdomen and male head. The top middle and right images are the result of a fusion of the number of FBP versions, performed with the trained ANN's. By visual impression, the noise-resolution balance in the fused images $\hat{f}$ is better than in any of the FBP versions feeding the process. The texture of the tissues is closer to the original (observed in the reference image, top left). The level of streaks and general noise are lower than in the central and right FBP images, and the image sharpness is higher than in the left and the central images. Thus, the fusion images enjoy the good properties offered by each of the FBP versions and are superior to all of them. The deeper ANN (top right) leads to better performance in SNR and SSIM (reported in Table I) but the differences are difficult to see by the naked eye.

The quantitative error measures we compute for this comparison include plain SNR, SNR windowed (SNRW) and the SSIM measures. These values are given in Table I. As can be seen, the fusion results' SNR is significantly higher, when compared with the best attainable FBP outcome. Specially, for the abdomen (male and female) and thorax sections, the improvement is above 1.75dB. The deeper ANN brings a further improvement of up to 1dB. Finally, the SSIM measure supports the above conclusions regarding the superiority of the fusion results.

We also compare two cases of output vectors produced by the ANN. In the bottom row of Figure 5, the image on the right is produced by a fusion process where a single pixel is recovered by the ANN for each input vector. The image on the left is produced by computing 3×3 pixel neighborhoods of each pixel and averaging the overlapping regions. The visual difference between the two is negligible, and the difference in SNR is 0.09dB in favor of the averaging approach. Judging from this (and other similar) tests, we conclude that forcing

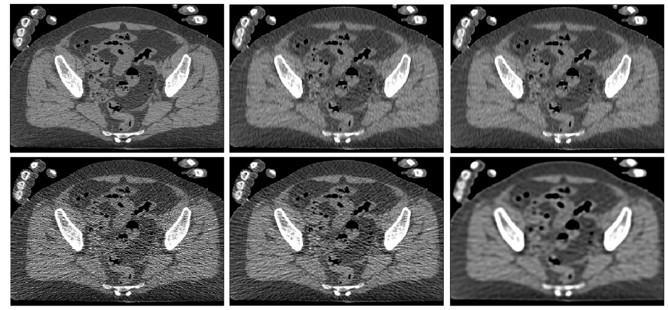

Figure 3: Female abdomen section. Top left: reference image. Top middle: the shallow ANN fusion result (SNR=25.36dB). Top right: the deep ANN fusion result (SNR=25.68dB). Bottom : FBP images participating in the fusion, produced with different low-pass filters (SNR=18.52, 20.1 and 23.11dB).

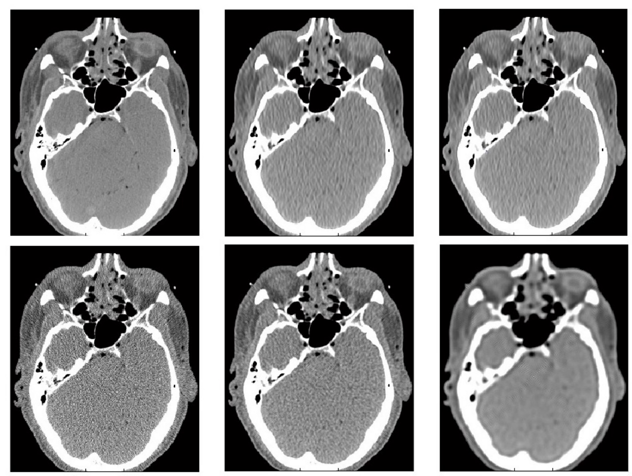

Figure 4: Male head section. Top left: reference image. Top middle: the shallow ANN fusion result (SNR=27.23dB). Top right: the deep ANN fusion result (SNR=27.35dB). Bottom: FBP images participating in the fusion, produced with different low-pass filters (SNR=22.89, 26.44 and 19.26dB).

Image	FBP $\phi_0 = 0.4$	FBP $\phi_0 = 2$	FBP $\phi_0 = \infty$	Fusion result	Fusion result (Deep ANN)
Abdomen (male)					
SNR	21.73	24.51	18.60	26.27	**27.27**
SNRW	21.91	23.90	19.28	26.81	**27.64**
SSIM	0.8060	0.7649	0.5308	0.8727	**0.9039**
Abdomen (female)					
SNR	20.10	23.11	18.52	25.36	**25.68**
SNRW	21.30	24.09	19.26	26.36	**26.50**
SSIM	0.8009	0.7952	0.5944	0.8727	**0.8803**
Thorax (male)					
SNR	21.66	23.97	19.05	25.74	**26.65**
SNRW	21.84	24.57	19.74	26.20	**27.01**
SSIM	0.8272	0.8141	0.6137	0.8831	**0.9068**
Head (male)					
SNR	19.26	26.44	22.89	27.23	**27.35**
SNRW	19.31	27.69	25.29	27.84	**28.24**
SSIM	0.9198	0.9307	0.8094	0.9419	**0.9522**

Table I: Quantitative measures for the FBP reconstructions and the fusion result. Best results are shown in bold.

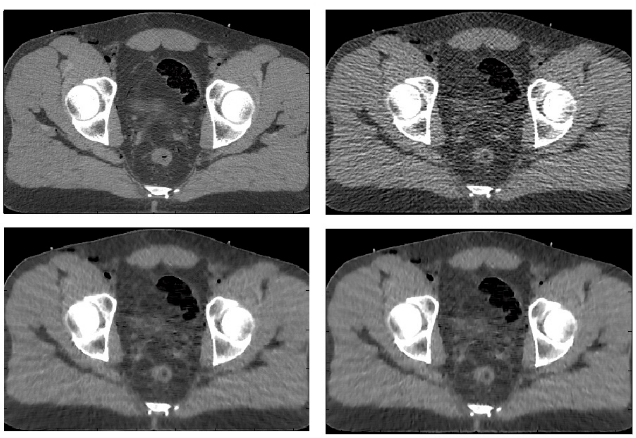

Figure 5: Thighs section. Top left: reference image. Top right: best-SNR FBP reconstruction (SNR=23.65dB). Bottom left: fusion result where ANN output size is 9 pixel (SNR=27.95dB). Bottom right: a fusion result where the ANN produces a single pixel value (SNR=27.86dB).

the neural networks to evaluate a number of pixels in the neighborhood of the one being recovered does not reduce its performance. Our extensive tests suggests that ANN with a single pixel output is sufficient, and does not lead to excessive artifacts.

B. Single-Image "Fusion"

A special case of the proposed method is to perform local processing with the ANN using only one FBP image. This, in fact, is a post-processing algorithm based on a regression function, which implements some non-linear local filter. In the following experiment we compare the performance of two

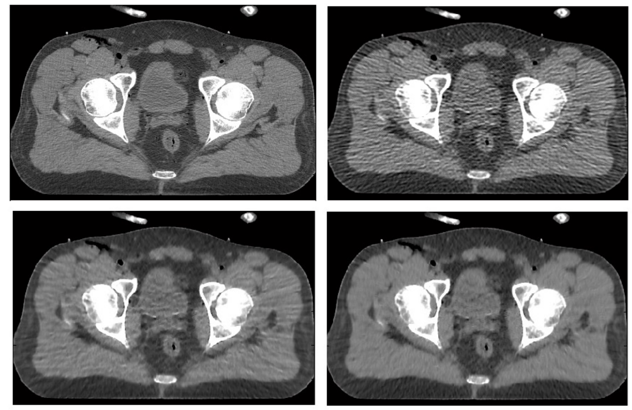

Figure 6: Thighs section. Top Left: reference image. Top Right: Best FBP version (SNR=25.25dB).Bottom Left: ANN fusion of a single FBP image with no sinogram filter, (SNR=28.35dB). Bottom Right: ANN fusion of three FBP versions, corresponding to filter cut-off frequency of $\phi_0 = [0.4, 2, \infty]$ (SNR=28.60dB).

fusion methods, one using three FBP images (sharp, normal and blurred) and another using only one FBP image produced with no low-pass filter. The results are displayed in Figure 6. Visually, in the single-image fusion some strong artificial streaks are observed, which do not appear in the multi-image fusion. Quantitatively, the SNR and the SSIM are also higher using the multi-image fusion. However, it is clear that in comparison to the best FBP version, even using the single image fusion leads to an improvement of more than 3dB in this example.

Since the single image processing done here may be interpreted as a post-processing stage on the FBP output, one may wonder if ”classical” image denoising algorithms could be of competition to the approach taken in this work. In Figure 7 we compare a single-image fusion result with a denoising on an FBP image using the well known BM3D. We have used the matlab code provided by the authors of [38], which is available at http://www.cs.tut.fi/ foi/GCF-BM3D/. BM3D algorithm was applied on several reconstructed FBP images corresponding to different filter cut-off frequency, only the best result is shown here. As the noise variance is one of the input to the BM3D algorithm, several values were tested and the optimal one was chosen. We can see that the single-image fusion outperforms the BM3D result both visually and in term of SNR. In the BM3D image, strong artifacts are observed and the image is oversmoothed. This conclusion was expectable as the BM3D algorithm is efficient in removing white and additive Gaussian noise while here the corruption process in much more complex and leads to colored noise as well as structured artifacts. On the other hand, the ANN is trained to remove precisely these kind of corruptions and indeed leads to significantly better results.

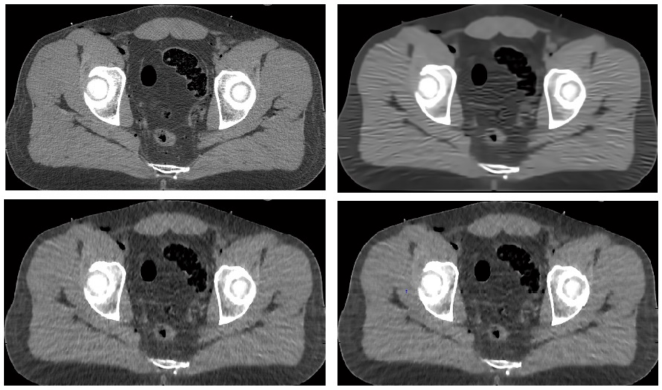

Figure 7: Thighs section. Top Left: reference image. Top Right: FBP image after BM3D denoising (SNR=27.56dB). Bottom Left: ANN fusion of a single FBP image with no sinogram filter (SNR=28.71dB).Bottom Right: Deep ANN fusion of a single FBP image with no sinogram filter (SNR=29.03dB).

VII. PWLS Boost - Algorithm Design and Empirical Study

A. Algorithm Description

The iterative PWLS algorithm (see Section II) can be boosted by gathering intermediate versions of the image at different numbers of iterations and fusing them using a trained ANN. We should emphasise that the proposed approach is one of post processing nature, implying that there is no feedback loop returning the neural network output back to the iterative reconstruction process. Introducing such feedback is indeed interesting, but outside the scope of this article. The idea is to capture the gradual transformation of the image from the initial to the final state. If the initial image is a blurred one, it gradually changes along the iterations towards a sharper version; the intermediate stages contain important information that can contribute to further improve the algorithm output. The fusion can also be performed over the parameter β – the weight of the penalty term (see Equation II.6), but this requires to run the iterative algorithm several times, which can be time consuming.

The method is very similar to the one proposed in the previous section. At the training stage, a CT reconstruction is performed with a high-quality reference at hand. The examples for ANN training are produced in the following manner: the vector of inputs, corresponding to a location q in the image, is assembled using neighborhoods of q in the different versions of the image, gathered along the PWLS iterations, or obtained using different penalty term weights. Specifically, we take a small neighborhood of pixels from each image in this sequence (see details below). The “correct answer”, corresponding to this vector of ANN inputs, is the value of the pixel q in the reference image. In the PWLS reconstruction the initial image could be a zero image or an image already reconstructed using the FBP algorithm, and both options are explored in the next section. The evolution of the SNR (measured on a male abdomen section) along the PWLS iterations is shown in Figure 8, the different graphs correspond to different penalty term weights β and the two possible initializations. Obviously, the initialization with zeros is simpler and less time consuming, and although both lead to very similar reconstruction results as can be seen in the graphs, the fusion results are better using the FBP initialization. Actually, the reconstruction process is quite different, in the first case it starts from all zeros and the image becomes sharper and sharper along the iterations while in the second case it starts from a noisy version and becomes smoother and smoother.

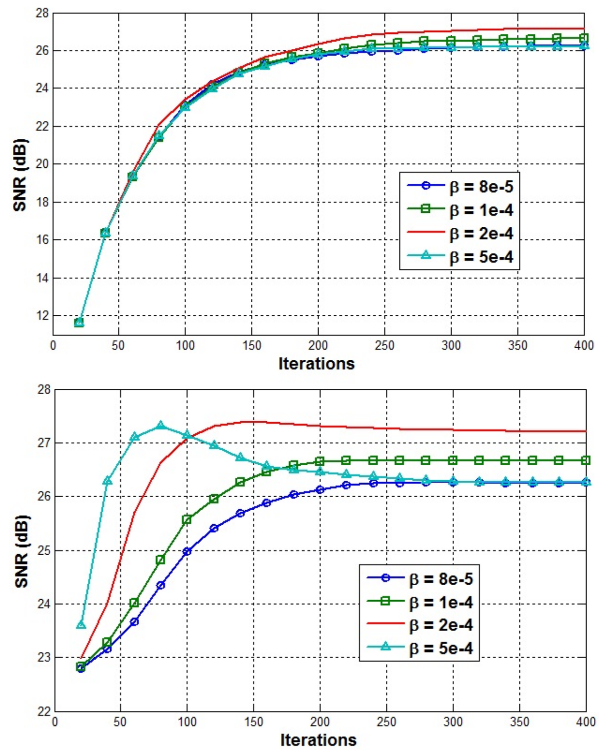

Figure 8: SNR evolution along PWLS iterations (measured on a male abdomen section). Top: Zero image initialization. Bottom: FBP image initialization.

B. PWLS Boost - Empirical Study

We conducted numerical experiments to demonstrate the proposed method using the same setup as in the FBP experiment. Training data for the ANN was obtained using a data-set of 12 axial male abdomen section images, and the model is tested on images from the male abdomen (10 cm away from the training region), thorax and head and the female abdomen. We will show results using zero image and FBP initializations. In the second case, an initial image $\tilde{f}$ is computed with the FBP algorithm using a sinogram filter with cut-off frequency value of 2.0 (see Figure 2). The PWLS algorithm is implemented as described in Section II, with $\delta = 0.02$. We have tested different values of the penalty term weights $\beta = [8e-5, 1e-4, 2e-4, 5e-4]$, and for each we have performed 400 PWLS iterations, saving an image version every 20 iterations – overall we have a sequence of 80 images. In practice, as done earlier in the context of the FBP fusion, we use three images out of this sequence, and the way to choose them is empirical – the goal is to collect image versions with different characteristics (smoother, noisier but sharper). We have tested several configurations and we report here on several interesting combinations. As mentioned previously, fusing images obtained using different penalty weights generally lead to better performance but this approach is more complex. Patches of size 7×7 are extracted from the three images and are used for the estimation of the output pixel neighborhood of size 3×3. Overall, the ANN has $3 \times 49 = 143$ inputs, and as for the FBP fusion, we have used different sizes and depths of ANN. We report the results obtained with two configurations - 3 and 7 hidden layers, each containing 20 neurons. These settings were obtained with a manual tuning of the design parameters.

In Figure 9 we display the fusion result obtained collecting versions along one PWLS reconstruction (using $\beta = 8e-5$) when the initial image was the FBP version. The fusion results have a higher visual quality than the three input images. Comparing to those images, the noise level in the fusion image is the lowest, and the tissue texture is closer to the original. The SNR values (stated in the Figure) also point to the improvement in quality. The SSIM of the fusion images are 0.9121 and 0.92 for shallow and Deep ANN respectively, while the sequence of PWLS results have the SSIM values of $0.7419, 0.8166, 0.8598$ (corresponding to the bottom row of Figure 9, left to right).

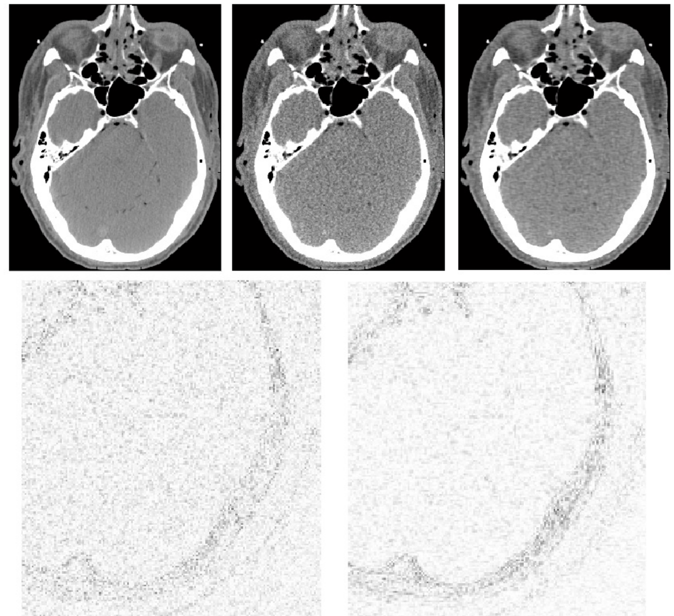

Figure 10: Male head section. Top Row: Left to right: reference image, the best PWLS version (SNR=26.72dB), ANN-fused PWLS with deep ANN (SNR=28.91dB). Bottom Row: absolute-value error images for the best SNR PWLS (left) and for the fusion image (right).

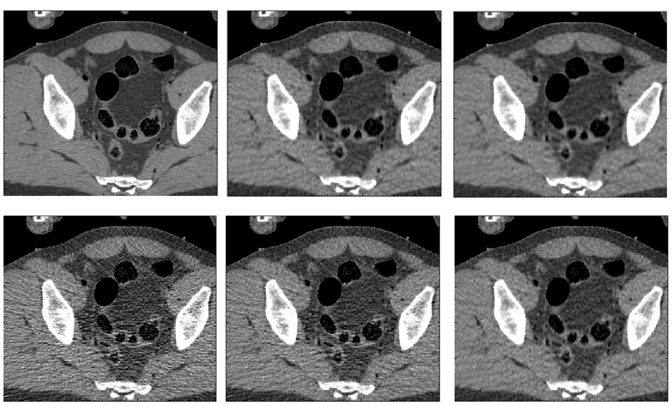

Figure 9: Images of abdomen section. Top row, left to right: reference image, ANN-fused PWLS with shallow ANN (SNR=28.20dB), ANN-fused PWLS with deep ANN (SNR=28.15dB). Bottom row: three PWLS versions (20 iterations, SNR=22.92, 100 iterations, SNR=25.07dB, 400 iterations, SNR=26.36dB).

A reconstruction of an additional test with zero initialization is displayed in Figure 10, the bottom part presents the absolute-value error images. The effect of the fusion observed here is similar to the one in the previous reconstruction, and in particular the noise effect and the texture are much closer to the original.

Table II reports the quantitative error measures for several different images. The first column shows the best PWLS results (after 400 iterations), the second column reports the fusion result when the PWLS algorithm was initialized with a zero image (images were collected after 160, 260 and 400 iterations), and the third column reports the fusion results when PWLS was initialized with an FBP version (images were collected after 20, 100 and 400 iterations). Both shallow and deep ANN results are reported. As can be seen, the fusion results always outperform the best PWLS reconstruction, with at least 0.8dB of improvement in SNR. The SSIM also support this conclusion. As mentioned previously, using an FBP initialization generally leads to better performance (true for all the images except from the head section). Obviously this approach is slightly more complex as the FBP construction needs to be computed before running the iterative process. If the PWLS algorithm can be ran several times with different penalty weights, fusing the different images (obtained after convergence) can lead to slightly better performance, but these improvements are of the order of few hundredth of dBs and are hard to see with the naked eye.

The ANN-based fusion can also contribute to the iterative reconstruction, without requiring to run it until convergence. Indeed, collecting images along the first iterations and fusing them with the presented model can shorten the reconstruction process significantly while preserving (or even improving) the results quality, as the fusion computional cost is lower by an order of magnitude than that of the iterative process. In Figure 11 the test is performed on another male thorax section, three versions are collected after 20, 40 and 60 iterations (Top images) and fused (Bottom right). The result is compared to the outcome of the whole iterative process after convergence (Bottom middle). In this example fusing images obtained

Image	PWLS 400 iters.	Fusion Result (init. with zeros) Shallow/Deep	Fusion Result (init. with FBP) Shallow/Deep
Abdomen section (male)			
SNR	26.36	27.69 / 28.40	28.15 / **28.60**
SNRW	27.17	28.22 / 28.83	28.63 / **28.95**
SSIM	0.8600	0.8999 / 0.9192	0.9121 / **0.923**
Abdomen section (female)			
SNR	25.19	26.07 / 26.52	26.35 / **26.63**
SNRW	26.19	26.76 / 27.11	27.03 / **27.25**
SSIM	0.8565	0.8870 / 0.8982	0.8939 / **0.8995**
Thorax section (male)			
SNR	25.64	26.98 / 27.55	27.71 / **27.88**
SNRW	26.41	27.48 / 27.96	28.10 / **28.23**
SSIM	0.8686	0.9047 / 0.9196	0.9210 / **0.9254**
Head section (male)			
SNR	26.72	28.5 / 28.91	28.15 / **29.02**
SNRW	28.94	29.75 / 29.81	28.99 / **30.12**
SSIM	0.919	0.9575 / 0.9692	0.9653 / **0.9723**

Table II: Quantitative measures for the PWLS reconstructions and the fusion results. Best results are shown in bold.

after only 60 iterations outperforms the result of the whole iterative process. Visually the fusion result is less noisy, the sharpness is quite similar, and obviousy the run time is reduced significantly. For this test the PWLS algorithm was initialized with an FBP version; initializating it with a zero image will naturally require more iterations to get a satisfying result.

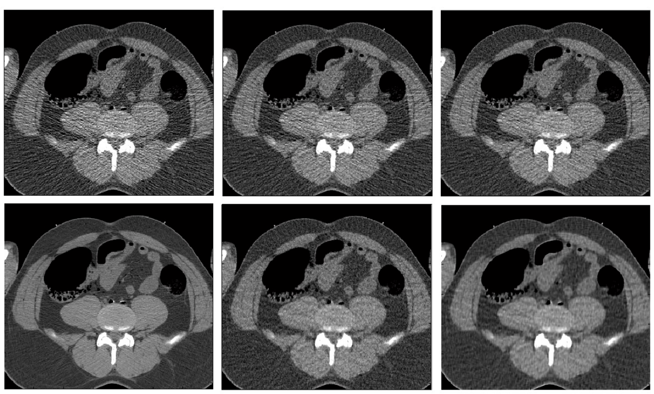

Figure 11: Top part: three PWLS versions (20, 40 and 60 iterations). Bottom part: original, PWLS after convergence (400 iterations) SNR = 26.40dB and fusion result SNR = 28.09dB.

C. Single-Image "Fusion"

As a last experiment, we consider the special case where the ANN only performs a local filtering of a single image. The fusion (in fact, post-processing) result is visually compared in Figure 12 versus the image produced by fusing 3 PWLS versions, as before. It can be observed that working with a single input reduces the noise appearing in the PWLS image, but it is slightly inferior to the fusion image produced from several PWLS versions. In table III, the results of fusion over 3 images obtained at early iterations (20, 40 and 60 iterations) is compared with single-image fusion for several images and with the result obtained after PWLS convergence (after 400 iterations). For all these images the fusion results outperform (both in terms of SNR and SSIM) the result obtained by the plain PWLS.

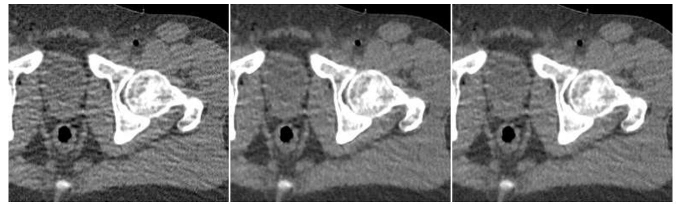

Figure 12: Left to right: PWLS image (60 iterations, SNR=26.02dB), single-image fusion (SNR=27.25dB), multi-image fusion (SNR=27.44dB).

Image	PWLS 20 / 40 / 60 iters.	PWLS 400 iters.	Fusion result Single / Full
Abdomen (male)			
SNR	22.92 / 23.26 / 23.77	26.36	27.61 / **28.03**
SSIM	0.7419 / 0.7539 / 0.7721	0.8599	0.9158 / **0.9193**
Abdomen (female)			
SNR	22.51 / 22.79 / 23.22	25.19	25.83 / **26.16**
SSIM	0.7752 / 0.7838 / 0.7967	0.8565	0.8915 / **0.8958**
Thorax (male)			
SNR	23.41 / 23.84 / 24.27	25.64	27.72 / **28.16**
SSIM	0.7971 / 0.8084 / 0.8206	0.8686	0.9358 / **0.9393**
Head (male)			
SNR	26.27 / 26.46 / 26.62	26.72	27.67 / **27.92**
SSIM	0.9176 / 0.9178 / 0.9180	0.919	0.9716 / **0.9721**

Table III: Quantitative measures for the PWLS reconstruction and fusion on early iterations. Best results are shown in bold.

VIII. Computational Complexity of The Method

In our experiments neural network with four hidden layers and 20 neurons in each layer, had about $k = 5600$ weights, *i.e.* we perform this number of summations and multiplications per reconstructed pixel. This is comparable to the computational cost of FBP or to a single iteration of PWLS reconstruction, when image size is 512×512.

In terms of actual run-time, we have measured that a single PWLS iteration takes 0.42sec on average, the FBP reconstruction requires 0.14sec and the application of the trained four hidden layers Neural Network is taking 0.21sec. All these were obtained using Matlab code on a 8-core @3.4 GHz PC with 16 GB RAM.

When our method is used with the FBP reconstruction, a number of FBP versions must be produced; in our experiments three reconstructions suffice. Therefore producing the fusion image will require roughly four times the extent of a single reconstruction (three FBP processes and the fusion step). Of course, the regular FBP image will be available for the radiologist after the usual time of a single FBP reconstruction.

As for the iterative PWLS algorithm, no changes in the reconstruction process are needed, since we only sample images along the standard iterations. The neural network fusion adds a cost comparable to one iteration, however, as we observe in the experiments, the high quality reconstruction is obtained after 60 iterations instead of 400, which saves the time significantly.

IX. Summary

We have introduced a method for quality improvement for a general parametrized signal estimator. The concept is to use a regression function for a local fusion of a number of estimator's outputs, corresponding to different parameter settings. The regression proposed is realized with feed-forward artificial neural networks. The fusion process consists of two components: first, the behavior of the signal in its different versions is gathered; second, the ANN performs its own non-linear filtering of the signal versions in small neighborhoods of the estimated pixel.

The proposed method is very general and CT reconstruction is only one possible application for it. The local fusion can be used to solve any linear on non-linear inverse problem where an algorithm, producing a solution estimate, exists. The proposed method will enable to incorporate the algorithm outputs, produced with different values of a core parameter, to a single improved result, thus removing the need for tuning this parameter.

In this work this concept was illustrated for the case of CT reconstruction, done with two basic algorithms – the FBP and the PWLS. Empirical results suggest that the local fusion can improve on the resolution variance trade-off of the given reconstruction algorithm, thus adding to the visual quality of the CT images. The post-processing method is not very time-consuming, and the cost of the local fusion can be well below the extent of one FBP reconstruction.

References

[1] J. Bian P.J. La Riviere and P.A. Vargas, "Penalized-likelihood sinogram restoration for computed tomography," *IEEE Trans. Med. Imag.*, vol. 25, no. 8, pp. 1022–1036, Aug. 2006.

[2] L. A. Shepp and J. B. Kruskal, "Computerized tomography - the new medical x-ray technology," *The American Mathematical Monthly*, vol. 85, no. 6, pp. 420–439, 1978.

[3] I.A. Elbakri and J.A. Fessler, "Statistical image reconstruction for polyenergetic x-ray computed tomography," *IEEE Trans. Med. Imag.*, vol. 21, no. 2, pp. 89–99, Feb. 2002.

[4] H. Lu Z. Liang J. Wang, T. Li, "Penalized weighted least-squares approach to sinogram noise reduction and image reconstruction for low-dose x-ray computed tomography," *IEEE Trans. Med. Imaging*, vol. 25, no. 10, pp. 1272–1283, 2006.

[5] Simon Haykin, *Neural Networks: A Comprehensive Foundation*, Macmillan, 1994.

[6] K. D. Sauer J. Hsieh J.-B. Thibault, C. A. Bouman, "A recursive filter for noise reduction in statistical iterative tomographic imaging," in *SPIE/IS&T Conference on* Computational Imaging IV. 2006, vol. 6065, pp. 60650X– 60650X–10, SPIE.

[7] Anja Borsdorf, "Adaptive filtering for noise reduction in x-ray computed tomography," in *Ph.D. Thesis*, 2009.

[8] K. M. Hanson, "Noise and contrast discrimination in ct," *Radiology of the Skull and Brain, Vol. V: Technical Aspects of Computed Tomography, T. H. Newton and D. G. Potts, eds.*, pp. 3941–3955, 1981.

[9] G. N. Ramachandran and A. V. Lakshminarayanan, "Three-dimensional Reconstruction from Radiographs and Electron Micrographs: Application of Convolutions instead of Fourier Transforms," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 68, no. 9, pp. 2236–2240, 1971.

[10] F. Wubbeling F. Natterer, *Mathematical methods in image reconstruction*, SIAM, 2001.

[11] Sathish Ramani and Jeffrey A Fessler, "A splitting-based iterative algorithm for accelerated statistical x-ray ct reconstruction," *Medical Imaging, IEEE Transactions on*, vol. 31, no. 3, pp. 677–688, 2012.

[12] J. Fessler, "Iterative methods for image reconstruction," in *ISBI Tutorial*. 2006, http://www.eecs.umich.edu/ fessler/papers/files/talk/08/isbi-notes.pdf.

[13] Stephen J. Nocedal, Jorge; Wright, *Numerical Optimization*, Springer-Verlag, 2-nd edition, 2006.

[14] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics Of Control, Signals, And Systems*, vol. 2, no. 4, pp. 303–314, 1989.

[15] G. E. Hinton D. E. Rumelhart and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, pp. 533 – 536, 1986.

[16] M. Egmont-Petersen, D. de Ridder, and H. Handels, "Image processing with neural networksa review," *Pattern Recognition*, vol. 35, no. 10, pp. 2279 – 2301, 2002.

[17] C. Nagar M. Cenci and A. Vecchione, "Papnet-assisted primary screening of conventional cervical smears," *Anticancer Res*, vol. 20, no. 5C, pp. 3887–3889, 2000.

[18] J.H.C.L. Hendriks G.M.T. Brake, N. Karssemeijer, "An automatic method to discriminate malignant masses from normal tissue in digital mammograms," *Phys. Med. Biol.*, vol. 45, no. 10, pp. 2843–2857, 2000.

[19] J.S. Lin K.S. Cheng and C.W. Mao, "The application of competitive hopfield neural network to medical image segmentation," *IEEE Trans. Med. Imag.*, vol. 15, no. 4, pp. 560–567, 1996.

[20] V.J. MartInez J.R. Hilera, M. Mazo, "Neural networks for ecg compression and classification," in *Proceedings of the 3rd International Conference on Fuzzy Logic, Neural Nets and Soft Computing*. 1994, pp. 121–124, Iizuka, Japan.

[21] S. Osowski and T. H. Linh, "Ecg beat recognition using fuzzy hybrid neural network," *IEEE Trans. On Biomedical Engineering*, vol. 48, no. 11, pp. 1265 – 1271, 2001.

[22] I. Aizenberg, N. Aizenberg, J. Hiltner, C. Moraga, and E. Meyer zu Bexten, "Cellular neural networks and computational intelligence in medical image processing," *Image and Vision Computing*, vol. 19, no. 4, pp. 177 – 183, 2001.

[23] Alexander M Bronstein, Michael M Bronstein, Michael Zibulevsky, and Yehoshua Y Zeevi, "Optimal nonlinear line-of-flight estimation in positron emission tomography," *Nuclear Science, IEEE Transactions on*, vol. 50, no. 3, pp. 421–426, 2003.

[24] R. Murugesan DC Durairaj, MC Krishna, "A neural network approach for image reconstruction in electron magnetic resonance tomography," *Comput. Biol. Med.*, vol. 37, no. 10, pp. 1492–501, 2007.

[25] Jr. C. E. Floyd, "An artificial neural network for spect image reconstruction," *IEEE Trans. Med. Imag.*, vol. 10, no. 3, pp. 485–487, 1991.

[26] C.E. Floyd Jr. and G.D. Tourassi, "An artificial neural network for lesion detection on single-photon emission computed tomographic images," *Investigative radiology*, vol. 27, no. 9, pp. 667–672, 1992.

[27] J. P. Kerr and E. B. Bartlett, "Medical image processing utilizing neural networks trained on a massively parallel computer," *Computers in Biology and Medicine*, vol. 25, no. 4, pp. 393–403, Jul 1995.

[28] E.B. Bartlett J.P. Kerr, "Neural network reconstruction of spect images," *J. Digital Imaging*, vol. 8, no. 3, pp. 116–126, 1995.

[29] R. Cierniak, "A 2d approach to tomographic image reconstruction using a hopfield-type neural network," *Artificial Intelligence in Medicine*, vol. 43, pp. 113–125, 2008.

[30] A. Adler and R. Guardo, "A neural network image reconstruction technique for electrical impedance tomography," *IEEE Trans. Med. Imag.*, vol. 13, no. 4, pp. 594–600, 1994.

[31] B.S. Hoyle A.Y. Nooralahiyan, "Three-component tomographic flow imaging using artificial neural network reconstruction," *Chemical Engineering Science*, vol. 52, no. 13, pp. 2139–2148, 1997.

[32] A. Goldenshluger and A. Nemirovsky, "On spatial adaptive estimation of nonparametric regression," *Math. Meth. Statist.*, vol. 6, no. 2, pp. 135–170, 1997.

[33] Yangqing Jia, Evan Shelhamer, Jeff Donahue, Sergey Karayev, Jonathan Long, Ross Girshick, Sergio Guadarrama, and Trevor Darrell, "Caffe: Convolutional architecture for fast feature embedding," *arXiv preprint arXiv:1408.5093*, 2014.

[34] L.Vese T.Chan, "Active contours without edges," *IEEE Trans. Im. Proc.*, vol. 10, no. 2, pp. 266–277, 2001.

[35] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr 2004.

[36] Yann LeCun, Léon Bottou, Genevieve B Orr, and Klaus-Robert Müller, "Efficient backprop," in *Neural networks: Tricks of the trade*, pp. 9–50. Springer, 1998.

[37] Ackerman M J, "The visible human project," *Proc. IEEE 86 504-11*, 1998.

[38] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian, "Bm3d image denoising with shape-adaptive principal component analysis," in *PROC. WORKSHOP ON SIGNAL PROCESSING WITH ADAPTIVE SPARSE STRUCTURED REPRESENTATIONS (SPARS09*, 2009.