

Image Processing via Calculus of Variations

Joaquín Mir Macías, Miguel Montes Lorenzo, Manuel Rodríguez Villegas

October 2025

Contents

1	Introduction	1
2	The Brachistochrone Problem	2
3	Image Processing	5
3.1	Rudin–Osher–Fatemi Model	5
4	Applied Example	6
5	Application in Medical Images	8
5.1	Classical Method (Rudin–Osher–Fatemi)	8
5.2	Method with Spatial Adaptation	9
5.3	Higher-Order Method (Regularizers)	10
6	Conclusion	10

1 Introduction

The calculus of variations constitutes one of the fundamental tools in applied mathematics when one wishes to optimize a quantity that depends on an unknown function. Unlike classical differential calculus, which focuses on maximizing or minimizing functions of a few real variables, in variational calculus the unknown is an entire function, and the objective consists of studying a functional that assigns a real value to each admissible function. This approach allows for the natural modeling of physical, geometric, or signal processing systems whose global behavior is described by energy integrals.

A historical example that illustrates this idea is the Brachistochrone problem, posed by Johann Bernoulli in 1696. The problem consists of determining which curve a particle travels along, subject only to gravity, when descending between two points in the minimum possible time. Although intuition might suggest that the solution is a straight segment, variational analysis leads to a differential equation whose solution is a cycloid. This formulation motivated the systematic development of the calculus of variations and fundamental tools such as the Euler–Lagrange equation and the Beltrami identity, which remain pillars of modern theory.

In recent decades, these same ideas have become especially relevant in image processing. Many classical tasks, such as noise reduction, edge recovery, segmentation, or *inpainting*, can be expressed through functionals that balance two complementary effects: fidelity to the observed data and the regularity or smoothness desired for the restored image. In this context, total variation (TV) has proven to be an especially effective regularizer, as it eliminates noise while preserving relevant discontinuities, something that does not occur with more traditional quadratic models.

The Rudin, Osher, and Fatemi (ROF) model, proposed in 1992, is one of the most influential examples of this approach. The model seeks a restored image by minimizing a functional that combines total variation with a fidelity term in the L^2 norm. The presence of the TV term allows for the preservation of sharp edges, while the regularization parameter controls the balance between smoothing and fidelity to the original image. Despite its simple formulation, the impact of the ROF model has been notable in theory and applications, giving rise to numerous extensions used today in computer vision and medical image processing.

The objective of this work is to show how the tools of the calculus of variations, introduced through the Brachistochrone problem, lead naturally to modern image restoration models. To do this, we begin by reviewing the statement of the classical problem and the derivation of the Euler-Lagrange equation. Next, we present the ROF model and analyze its geometric interpretation, its fundamental properties, and the effect of its parameters. Finally, we apply different methods based on total variation—both the classical model and recent extensions—to synthetic and medical images, allowing us to compare their behavior and discuss the advantages and disadvantages of each approach.

2 The Brachistochrone Problem

The formulation and subsequent resolution of the Brachistochrone problem marked the beginning of the calculus of variations, where one seeks to find relative extrema of continuous functionals defined over some function space.

Given points A and B in the vertical plane, the Brachistochrone problem consists of finding the curve followed by a particle moving from A to B in the minimum possible time, solely under the effect of gravity.

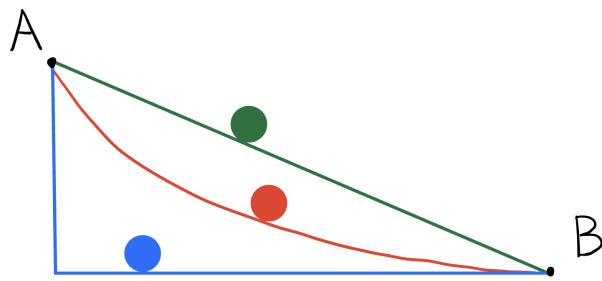


Figure 1: Possible solutions to the Brachistochrone problem.

It was posed by Johann Bernoulli in 1696 and solved by Newton in 1697, although five other mathematicians (Johann and Jakob Bernoulli, Leibniz, L'Hôpital, and Tschirnhaus) also intervened in the solution.

Problem Statement

To find the solution, we must minimize the time employed by the particle. To do this, we represent time as a function of the distance traveled and the velocity of the particle.

If the curve describing the trajectory is expressed as $y = y(x)$, the arc length between x_a and x_b is

$$s = \int_{x_a}^{x_b} \sqrt{1 + (y'(x))^2} dx,$$

with the notation $y' = \frac{dy}{dx}$.

Starting from the principle of conservation of energy, we know that kinetic energy and potential energy must be equal at all times. Therefore, if y is the vertical height measured from the starting point, the energy of the system is given by

$$\frac{1}{2}mv^2 = mgy \implies v = \sqrt{2gy},$$

where g is the acceleration of gravity and v is the velocity of the body at the point considered.

The time it takes for the particle to move between two points is given by $t = \frac{s}{v}$. To use the integral expression of the distance traveled, we can express the differential time as $dt = \frac{ds}{v}$ and obtain the total time

by integrating. Substituting ds and v , we obtain the time functional:

$$t = \mathcal{T}[y] = \int_{s_a}^{s_b} \frac{ds}{v} = \int_{x_a}^{x_b} \frac{\sqrt{1+y'(x)^2}}{\sqrt{2g}\sqrt{y(x)}} dx = \int_{x_a}^{x_b} f(y(x), y'(x)) dx,$$

where we have defined

$$f(y, y') := \frac{\sqrt{1+y'^2}}{\sqrt{2gy}}.$$

Given that we seek an extremum of the functional, the function $y(x)$ must satisfy the Euler-Lagrange equation:

$$\frac{\partial f}{\partial y} - \frac{d}{dx} \left(\frac{\partial f}{\partial y'} \right) = 0.$$

We observe that the function $f(y, y')$ does not depend explicitly on x (that is, $\frac{\partial f}{\partial x} = 0$). This allows us to find a first integral of motion.

We calculate the total derivative of f with respect to x using the chain rule:

$$\frac{df}{dx} = \frac{\partial f}{\partial y} y' + \frac{\partial f}{\partial y'} y'' + \frac{\partial f}{\partial x}.$$

As $\frac{\partial f}{\partial x} = 0$, we are left with:

$$\frac{df}{dx} = y' \frac{\partial f}{\partial y} + y'' \frac{\partial f}{\partial y'}.$$

From the Euler-Lagrange equation, we solve for $\frac{\partial f}{\partial y}$:

$$\frac{\partial f}{\partial y} = \frac{d}{dx} \left(\frac{\partial f}{\partial y'} \right).$$

Substituting this into the total derivative expression:

$$\frac{df}{dx} = y' \left[\frac{d}{dx} \left(\frac{\partial f}{\partial y'} \right) \right] + y'' \frac{\partial f}{\partial y'}.$$

We recognize that the right side is the derivative of a product $\frac{d}{dx}(uv) = u'v + uv'$:

$$\frac{df}{dx} = \frac{d}{dx} \left(y' \frac{\partial f}{\partial y'} \right).$$

Regrouping terms:

$$\frac{d}{dx} \left(f - y' \frac{\partial f}{\partial y'} \right) = 0.$$

This implies that the term in parentheses is constant. This is the so-called Beltrami Identity:

$$f - y' \frac{\partial f}{\partial y'} = C.$$

We calculate the necessary derivatives:

$$f(y, y') = \frac{\sqrt{1+y'^2}}{\sqrt{2gy}}, \quad \frac{\partial f}{\partial y'} = \frac{y'}{\sqrt{2gy}\sqrt{1+y'^2}}$$

Substituting, we arrive at:

$$\frac{\sqrt{1+y'^2}}{\sqrt{2gy}} - y' \frac{y'}{\sqrt{2gy}\sqrt{1+y'^2}} = \frac{1}{\sqrt{2gy}\sqrt{1+y'^2}} = C,$$

Solving for y' :

$$1 + y'^2 = \frac{1}{2gC^2y} \implies \frac{dy}{dx} = y' = \pm \sqrt{\frac{1}{2gC^2y} - 1}.$$

Parametric solution: Cycloid

The solution that satisfies $y(0) = 0$ can be expressed via angular parameters using the expression of the cycloid:

$$\begin{cases} x(\theta) = \frac{1}{4gC^2}(\theta - \sin \theta), \\ y(\theta) = \frac{1}{4gC^2}(1 - \cos \theta), \end{cases} \quad \theta \in [0, \theta_f].$$

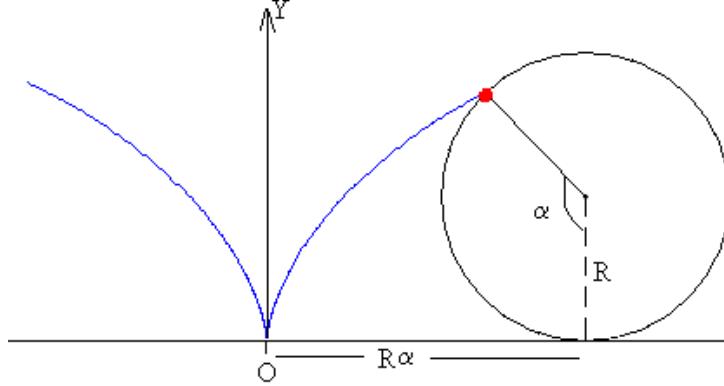


Figure 2: Construction of the cycloid curve.

We verify that this curve satisfies the equation for y' . Calculating derivatives with respect to θ :

$$\frac{dx}{d\theta} = \frac{1}{4gC^2}(1 - \cos \theta), \quad \frac{dy}{d\theta} = \frac{1}{4gC^2} \sin \theta,$$

therefore

$$\frac{dy}{dx} = \frac{\frac{dy}{d\theta}}{\frac{dx}{d\theta}} = \frac{\sin \theta}{1 - \cos \theta} = \sqrt{\frac{1 + \cos \theta}{1 - \cos \theta}}$$

If we substitute

$$y(\theta) = \frac{1}{4gC^2}(1 - \cos \theta)$$

into the expression

$$\sqrt{\frac{1}{2gC^2y} - 1},$$

we get:

$$\frac{1}{2gC^2y} = \frac{1}{2gC^2} \cdot \frac{1}{1 - \cos \theta} = \frac{1}{\frac{4gC^2}{2}} = \frac{2}{1 - \cos \theta}$$

Subtracting 1:

$$\frac{1}{2gC^2y} - 1 = \frac{2}{1 - \cos \theta} - 1 = \frac{2 - (1 - \cos \theta)}{1 - \cos \theta} = \frac{1 + \cos \theta}{1 - \cos \theta}.$$

Taking the square root:

$$\sqrt{\frac{1}{2gC^2y} - 1} = \sqrt{\frac{1 + \cos \theta}{1 - \cos \theta}}.$$

Therefore

$$\frac{dy}{dx} = \frac{\sin \theta}{1 - \cos \theta} = \sqrt{\frac{1 + \cos \theta}{1 - \cos \theta}} = \sqrt{\frac{1}{2gC^2y} - 1},$$

and the parametrization by the cycloid satisfies the condition imposed by the first integral.

3 Image Processing

Having analyzed the theoretical development of a variational calculus problem, let us delve into one of the multiple real-life applications: image processing.

3.1 Rudin–Osher–Fatemi Model

Given a noisy image $f(x, y)$, we wish to obtain a smoother version $u(x, y)$ that retains great similarity to the original image but removes much of the noise present. The model proposed by Rudin, Osher, and Fatemi [2] (1992) is based on minimizing the following functional:

$$J[u] = \frac{1}{2} \int_{\Omega} (u - f)^2 dx dy + \lambda \int_{\Omega} |\nabla u| dx dy, \quad (1)$$

where the first term measures data fidelity and the second, total variation (TV), acts as a regularization term.

The fidelity term helps the new image remain close to the original image f . The total variation term penalizes the magnitude of the gradient, controlling the roughness of the image without excessively punishing discontinuities.

Unlike quadratic regularization $\int |\nabla u|^2$, the linear penalty on $|\nabla u|$ allows for the existence of finite jumps in u . This causes high-frequency oscillations, normally related to noise, to be suppressed, and significant edges to be preserved.

The parameter $\lambda > 0$ controls the balance between smoothing and fidelity: large values of λ produce a very smooth image, removing much of the noise but significantly blurring the image (trying to make $|\nabla u|$ as low as possible), while small values leave more noise. As we will see later, its optimal adjustment depends on the noise level in the image.

This method also entails a series of limitations, among which stand out the loss of contrast in small homogeneous objects or *staircasing* artifacts (flat zones separated by jumps).

Derivation of the Euler–Lagrange condition

Consider the functional (1). The stationarity condition is obtained by imposing that the variation of $J[u]$ vanishes for every variation function η such that $\eta(a) = \eta(b) = 0$:

$$\frac{d}{d\varepsilon} J[u + \varepsilon\eta] \Big|_{\varepsilon=0} = 0.$$

The fidelity term becomes:

$$\frac{d}{d\varepsilon} \Big|_{\varepsilon=0} \frac{1}{2} \int (u + \varepsilon\eta - f)^2 dx dy = \int (u - f) \eta dx dy.$$

On the other hand, assuming $\nabla u \neq 0$ in the region of interest, the total variation term becomes:

$$\frac{d}{d\varepsilon} \Big|_{\varepsilon=0} \int |\nabla(u + \varepsilon\eta)| = \int \frac{\nabla u}{|\nabla u|} \cdot \nabla \eta dx dy.$$

Applying integration by parts and discarding the boundary term due to boundary conditions, we obtain

$$\int \frac{\nabla u}{|\nabla u|} \cdot \nabla \eta = - \int \nabla \left(\frac{\nabla u}{|\nabla u|} \right) \eta dx dy.$$

Summing both terms and requiring that the total variation be null for all η , we obtain

$$\int \left[(u - f) - \lambda \nabla \left(\frac{\nabla u}{|\nabla u|} \right) \right] \eta dx dy = 0.$$

Due to the arbitrariness of η , the Euler–Lagrange condition is satisfied:

$$(u - f) - \lambda \nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) = 0.$$

(2)

This is the fundamental equation of the Rudin–Osher–Fatemi model in its classical formulation.

Geometric interpretation of the curvature term

The vector

$$\frac{\nabla u}{|\nabla u|}$$

represents the unit normal direction to the level curves of u . Its divergence,

$$\nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right),$$

corresponds to the mean curvature of said curves in two dimensions.

In equation (2), this term causes a smoothing effect: it reduces the curvature of small contours (removing fine noise) but does not force the elimination of large discontinuities, which allows sharp edges to be preserved.

In homogeneous zones, the gradient is small and the equation diffuses intensity, averaging values and removing noise. At edges, where $|\nabla u|$ is large, the curvature term is attenuated and prevents transversal diffusion, thus preserving contours. The key is the normalization factor, which makes the curvature term more significant than one might expect with very small gradients and, in turn, normalizes it with very large gradients, preventing smoothing from crossing edges.

In practice, when $|\nabla u|$ is very small, the quotient $\nabla u/|\nabla u|$ becomes unstable. To avoid singularities, regularization is used:

$$|\nabla u| \approx \sqrt{|\nabla u|^2 + \varepsilon^2}, \quad \varepsilon \ll 1.$$

This guarantees numerical stability and allows solving (2) via iterative schemes.

A practical way to obtain u consists of performing a gradient descent of the functional $J[u]$ in an artificial time t :

$$\frac{\partial u}{\partial t} = \nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) + \lambda(f - u).$$

The steady state (when $\partial_t u = 0$) coincides with the Euler–Lagrange equation (2). This equation can be interpreted as a nonlinear diffusion: it diffuses within homogeneous regions and stops at edges.

In finite discretization, this flow produces a sequence u^k that converges towards the minimizer u_λ . In practice, a finite difference scheme is used, or the Chambolle dual projection method, widely adopted for its stability and speed.

Common extensions include:

- Anisotropic models, where diffusion is weighted in preferred directions.
- ROF with L^1 fidelity, more robust against impulsive noise (random white and black points).
- Second-order models (TGV, Total Generalized Variation) that reduce the "staircasing" effect (when stepped lines appear at image edges).

4 Applied Example

In this section, we explore in depth the power of this method in a daily use case.

We begin with a grayscale image f with value range in $[0, 255]$ contaminated with additive Gaussian noise. Below are the qualitative results of the ROF model with different values of λ :

- $\lambda = 0.8$: the image u_λ is very similar to f ; noise removal is slight.
- $\lambda = 0.05$: noise is significantly reduced, maintaining contours and essential structures.
- $\lambda = 0.01$: fine details disappear, obtaining a *cartoonized* image, where only global structures remain.



Figure 3: Result of the ROF model applied to the *Lena* image. Left: original noisy image f . Right: result u_λ for $\lambda = 0.05$.

Denoising Comparison: Gaussian vs. TV



Figure 4: Comparison with a Gaussian filter method.

Secondly, we compare the performance of the Rudin-Osher-Fatemi method, using its Python implementation in `tv_chambolle`, with a Gaussian noise filter in a denoising task.

Observing the image in Figure 4, we realize the main differences between both methods. The Gaussian filter, firstly, provides a color closer to the original but at the same time allows much noise to remain in the image. On the other hand, the total variation method manages to remove noise much more significantly in exchange for losing some color, probably due to the balance between fidelity and gradient limitation. In both cases, parameter values that provide the most visually robust result have been used.

The sensitivity analysis of the parameter λ in Figure 5 helps us better understand the implications of different values and when it is interesting to use one or the other.

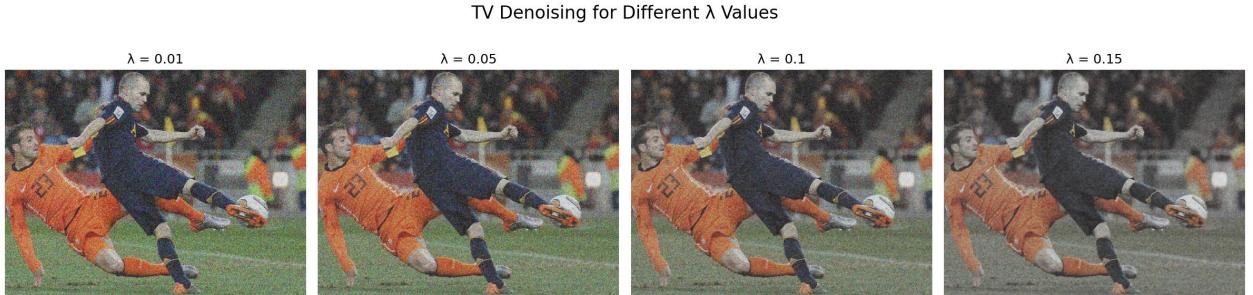


Figure 5: Sensitivity analysis of the parameter λ .

Thus, we clearly observe the importance of λ on the color of the resulting image. While a very low value will preserve color very effectively, it will also cause most of the noise to remain, as the gradient term is highly penalized and practically the entire minimization signal comes from fidelity to the original image.

By increasing λ and giving more importance to gradient limitation, we see that noise is much more blurred, but we end up losing so much color that it may only be worth it if we work with black and white images.

5 Application in Medical Images

Below is one of the most interesting applications of image processing: noise reduction in medical images. On many occasions, medical scanners present a high level of noise due to the difficult process of obtaining said data, usually through X-rays. For this reason, noise removal becomes a fundamental step prior to analysis by a doctor or a computer vision system [3].

In this section, we apply the noise reduction method explained previously to brain and lung scans. Furthermore, we compare three variants and discuss in which cases it is more convenient to use one or the other.

5.1 Classical Method (Rudin-Osher-Fatemi)

First, we employ the original Rudin-Osher-Fatemi method on medical scans. Depending on the parameter λ , we will obtain a result with more noise but better-preserved edges, or reduced noise at the cost of losing some definition.

Thus, we distinguish three cases in which the optimal parameter λ may be different.

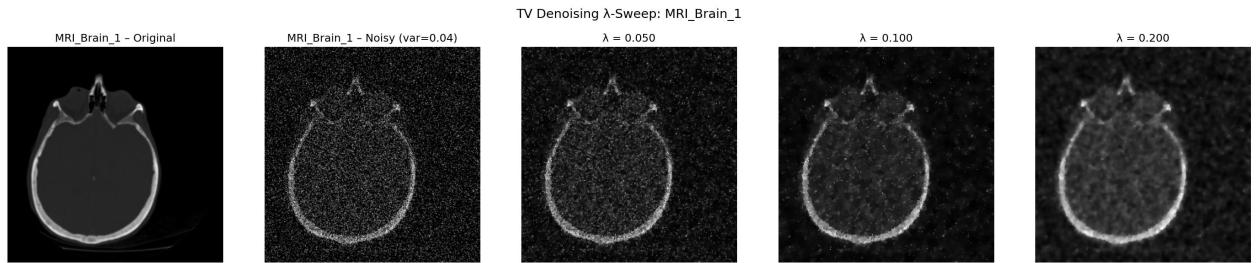


Figure 6: Brain scan.

In the brain scan of Figure 6, where there is a clear distinction between tissues (gray matter and white matter), we observe clearly how increasing λ greatly reduces image definition, so it is preferable to use a lower value that, despite maintaining some more noise, offers a result with higher resolution. The brain scan images have been obtained from the Python `skimage` package [4].

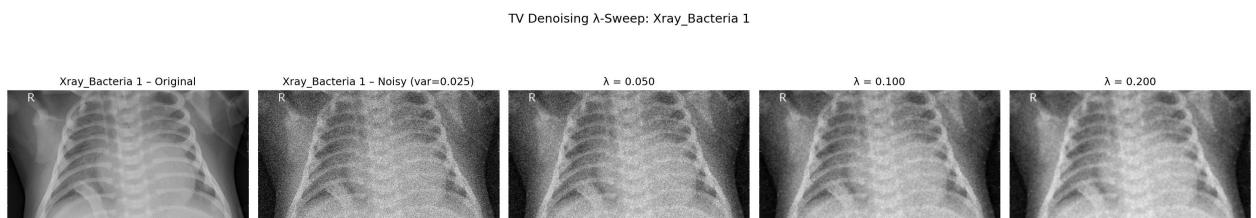
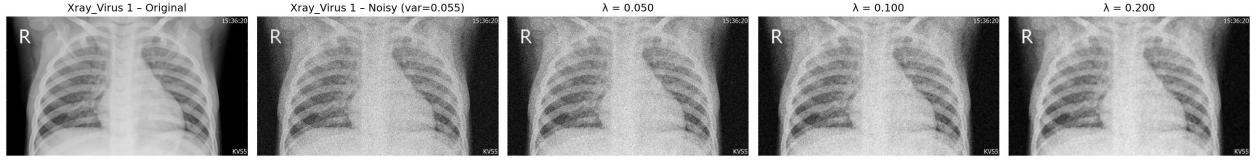


Figure 7: Lung scan in the presence of a bacterium.

In the lung scan of Figure 7, we observe how a smaller λ has almost zero influence on noise removal, while higher values do manage to remove it more satisfactorily. In this case, the loss of resolution may compensate for the superior noise removal, managing to obtain an image almost identical to the original.

Finally, observing the third image (Figure 8), we perceive an effect similar to the previous one, which is that in this type of scan it is usually preferable to lose some resolution to obtain an image almost completely free of noise. The lung scans have been obtained from the public Kaggle Chest X-Ray Images (Pneumonia) dataset [1].

**Figure 8:** Lung scan in the presence of a virus.

5.2 Method with Spatial Adaptation

To reduce noise without degrading fine contours in medical images, we consider an extension of the ROF model with *spatial weighting*. We define the functional

$$\min_u \left\{ \underbrace{\int_{\Omega} w(x) |\nabla u(x)| dx}_{\text{weighted TV}} + \underbrace{\frac{1}{2} \int_{\Omega} \lambda(x) (u(x) - f(x))^2 dx}_{\text{spatial fidelity}} \right\},$$

where $w : \Omega \rightarrow [0, 1]$ attenuates regularization near edges and $\lambda : \Omega \rightarrow \mathbb{R}_+$ allows adjusting data fidelity locally. The typical choice is constant $\lambda(x) \equiv \lambda > 0$ and a w dependent on a stable *edge indicator*:

$$w(x) = \exp\left(-\left(\frac{|\nabla(G_\sigma * f)(x)|}{k}\right)^\beta\right), \quad \beta \in [1, 4],$$

where G_σ is a Gaussian filter (robustness), $k > 0$ scales the edge threshold (e.g., k taken as a percentile of the gradient modulus), and $|\nabla(G_\sigma * f)|$ is the gradient modulus of the smoothed image.

Using the Euler-Lagrange equation, for sufficiently regular u and $|\nabla u| \neq 0$ a.e., the subgradient of the weighted TV is

$$\partial\left(\int w |\nabla u|\right) = -\nabla \cdot \left(w \frac{\nabla u}{|\nabla u|}\right).$$

The stationarity condition of the previous functional with Neumann boundary conditions ($\nabla u \cdot \nu = 0$ on $\partial\Omega$) leads to

$$0 = -\nabla \cdot \left(w \frac{\nabla u}{|\nabla u|}\right) + \lambda(x) (u - f),$$

or explicitly,

$$\boxed{\lambda(x) (u - f) - \nabla \cdot \left(w(x) \frac{\nabla u}{|\nabla u|}\right) = 0.} \quad (3)$$

We resolve (3) via gradient descent:

$$\partial_t u = \nabla \cdot \left(w \frac{\nabla u}{\sqrt{|\nabla u|^2 + \varepsilon^2}}\right) + \lambda(x) (f - u), \quad u(\cdot, 0) = f,$$

where $\varepsilon \ll 1$ stabilizes the norm and preserves strong edges. We discretize with finite differences and Neumann conditions, using a time step $\Delta t \in [0.1, 0.25]$ and a fixed number of iterations (200–300) until visual convergence.

To obtain the weights w , we start by calculating $g_\sigma = G_\sigma * f$, its gradient by finite differences, and the modulus $|\nabla g_\sigma|$. We take k as the percentile p of $|\nabla g_\sigma|$ to adapt the threshold to each image, and fix $\beta \in [1, 4]$. Thus, $w \approx 1$ in almost flat zones (more smoothing) and $w \approx 0$ along contours (smoothing inhibited), which maintains key anatomical structures. Figure 9 shows an example of this application.

Thus, as practical observations we can highlight:

- In brain MRI images, $w(x)$ preserves white/gray matter interfaces; moderate λ avoids oversmoothing clinically relevant texture.
- In chest X-rays, attenuation in fissures and lung edges allows for higher λ values to clean noise without losing fine anatomy.
- The choice of (σ, p, β) governs edge *selectivity*.

CT color + pesos espaciales + TV adaptativo canal a canal

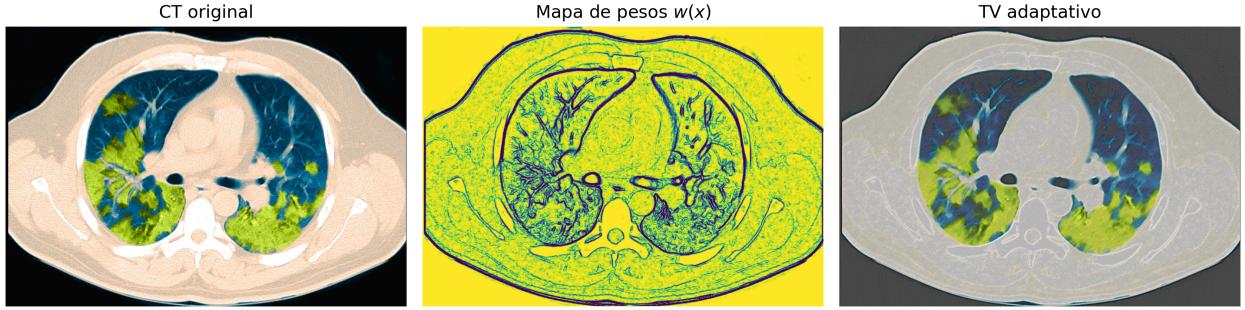


Figure 9: Spatial Adaptation in CT image

5.3 Higher-Order Method (Regularizers)

Finally, we study a denoising method that introduces higher-order regularization. The idea is simple: instead of penalizing only the image gradient, as occurs in ROF, we also add a term that penalizes the Laplacian. In this way, the model tends to produce smoother transitions and avoids the "staircasing" effect, which appears quite frequently when classical total variation is used.

The functional we want to minimize is

$$J[u] = \frac{1}{2} \int_{\Omega} (u - f)^2 dx + \frac{\alpha}{2} \int_{\Omega} |\nabla u|^2 dx + \frac{\beta}{2} \int_{\Omega} (\Delta u)^2 dx,$$

where f is the noisy image and the parameters α and β control the smoothing intensity. The term with α behaves like a standard first-order regularization while β controls the second-order effect, which allows removing abrupt changes without generating constant piecewise regions.

Deriving this functional yields the associated Euler–Lagrange equation:

$$0 = (u - f) - \alpha \Delta u + \beta \Delta^2 u,$$

where $\Delta^2 u$ is the bilaplacian. Solving this equation directly is not simple, so we use explicit gradient descent:

$$\partial_t u = (f - u) + \alpha \Delta u - \beta \Delta^2 u, \quad u(\cdot, 0) = f.$$

To implement this method we use finite differences, both for the Laplacian and for the bilaplacian (which we calculate by applying the Laplacian twice). The boundary conditions are Neumann to maintain the image within the expected range. In practice, it suffices to choose a small time step and perform several iterations until the changes between steps are minimal.

From a practical point of view, the method's behavior is somewhat better than the previous two. The higher-order regularizer generates much smoother images without artifacts, and usually removes noise quite uniformly. As a consequence, edges become somewhat less sharp, although without completely losing the general structure. This is noticeable above all in X-rays, where fine details disappear sooner than in the case of adaptive TV. In contrast, in brain scans, which have broad regions of almost constant intensity, the method works especially well because it reduces noise without introducing artificial patterns.

As an example, in Figure 10 we show a comparison of the method with different values of α and β . In all cases, a more homogeneous smoothing is observed compared to the ROF model, although at the cost of losing some detail in the edges.

6 Conclusion

This work has traced a path from the historical foundations of the calculus of variations to its contemporary application in image processing via Total Variation. It has been demonstrated how classical tools allow optimizing functionals both to minimize descent time in a mechanical problem and to reduce noise in a digital image. The analysis of the Rudin-Osher-Fatemi model has been central, showing its ability to balance fidelity to observed data and the regularity of the restored image.

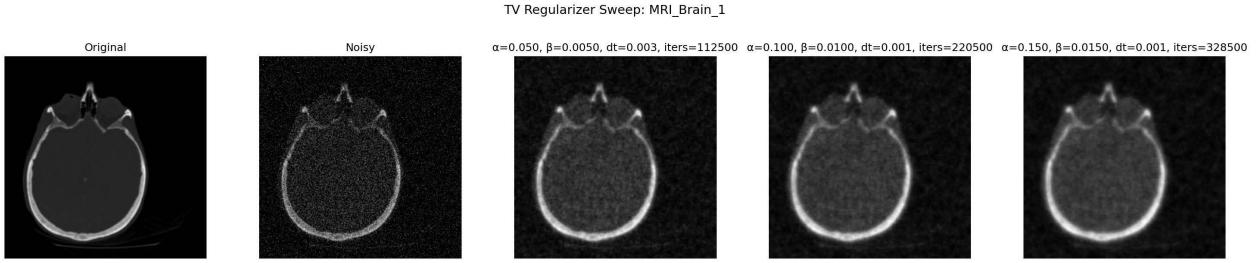


Figure 10: Results of the higher-order regularizer on a brain scan.

The practical examples, applied to both general images and medical scans, have validated the properties of the ROF model. The sensitivity analysis of the parameter λ has underscored the compromise between noise elimination and detail preservation. It has been found that lower λ values are preferable in images with fine structures, such as brain MRIs, while higher values may be necessary in X-rays to compensate for high noise levels. The comparison with Gaussian filters has also reinforced the superiority of TV for removing noise without excessively blurring contours.

Finally, the study has extended to more advanced methods to overcome the limitations of the classical ROF model. A model with spatial adaptation and a higher-order regularizer have been presented, producing smoother and more homogeneous results, especially useful in images such as magnetic resonance imaging. This comparative analysis demonstrates that variational methods offer a powerful and flexible set of tools, fundamental for current challenges in medical image analysis and computer vision.

The code used for the experiments can be found at https://github.com/rdgzmanuel/image_processing.

References

- [1] Paultimothy Mooney. Chest x-ray images (pneumonia) dataset. Kaggle dataset, 2018. Accessed: 2025-11-07.
- [2] Leonid I. Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1):259–268, 1992.
- [3] Francesco Tonolini, Jack Radford, Alex Turpin, Daniele Faccio, and Roderick Murray-Smith. Variational inference for computational imaging inverse problems, 2020.
- [4] Stéfan van der Walt, Johannes L. Schönberger, Juan Nunez-Iglesias, François Boulogne, Joshua D. Warner, Neil Yager, Emmanuelle Gouillart, Tony Yu, and the scikit-image contributors. scikit-image: Image processing in python. *PeerJ*, 2:e453, 2014.