Unsupervised image segmentation via maximum \textit{a posteriori} estimation of continuous max-flow

Ashif Sikandar Iquebal\textsuperscript{1} and Satish Bukkapatnam\textsuperscript{11}

\textsuperscript{1}Department of Industrial and Systems Engineering, Texas A&M University

Abstract—Recent thrust in imaging capabilities in medical as well as emerging areas of manufacturing systems creates unique opportunities and challenges for on-the-fly, unsupervised estimation of anomalies and other regions of interest. With the ever-growing image database, it is remarkably costly to create annotations and atlases associated with different combinations of imaging capabilities and regions of interest. To address this issue, we present an unsupervised learning approach to a continuous max-flow problem. We show that the maximum \textit{a posteriori} estimation of the image labels can be formulated as a capacitated max-flow problem over a continuous domain with unknown flow capacities. The flow capacities are then iteratively obtained by considering a Markov random field prior over the neighborhood structure in the image. We also present results to establish the consistency of the proposed approach. We establish the performance of our approach on two real-world datasets including, brain tumor segmentation and defect identification in additively manufactured surfaces as gathered from electron microscopic images. We also present an exhaustive comparison with other state-of-the-art supervised as well as unsupervised algorithms. Results suggest that the method is able to perform almost comparable to other supervised approaches, but more 90% improvement in terms of Dice score as compared to other unsupervised methods.

Index Terms—Continuous max-flow, Maximum likelihood estimation

1 INTRODUCTION

The goal of image segmentation is to cluster pixels of an image into logical entities/objects based on their intensity values and the neighborhood structure \cite{1}. Towards this, mainly two lines of thought have existed—graph-based approaches \cite{2}, \cite{3} and Gaussian mixture models (GMM) \cite{4}, \cite{5}, \cite{9}. GMM are typically sensitivity to noise, as they depend mostly on the distribution of pixel intensities. In contrast, graph-based approaches are region based, i.e., the resulting segmentation is influenced by the neighborhood structure, and guarantees optimality in terms of the energy functions (e.g., see Eq. (1) below) \cite{6}, \cite{7}. More recently, the graph-based approaches has been popularized in various fields, including medical image segmentation, object tracking in static and video surveillance and many more \cite{2}, \cite{8}, \cite{9}.

In general, graph-based approaches can be formulated as an energy minimization functional over a discrete image grid \cite{10} of the form:

\begin{equation}
E(l) = E_{\text{data}}(l) + E_{\text{smooth}}(l)
\end{equation}

where $E_{\text{data}}(l)$ accounts for the measure of fit of the labels $l = \{l_i, \ i = 1, 2, \ldots, n\}$, $n$ is the number of segments, and $E_{\text{smooth}}(l)$ controls the smoothness of the segmentation. Note that the discretization of the image domain typically results in metrication error due to the grid bias \cite{11}. To overcome this, we consider the image domain to be continuous, i.e., a bounded continuum of locations in $\mathbb{R}^2$, originally presented by Strang \cite{12}. However, a major short-comings is that the graph-based approaches are mostly supervised or require manual parameter settings (e.g., the flow capacities in the min-cut/max-flow problem \cite{13}).

In most graph cut algorithms \cite{14}, \cite{15}, \cite{16}, this is generally resolved by requiring the users to define the object of interest and background (e.g., by selecting sample pixels in background and foreground) to train the model. Apart from this, they use annotated atlases (reference templates) or training data, especially for the segmentation of anatomical structures in ever increasing medical images. Given the atlas, an image can be segmented by defining “anatomically correct” mapping from the coordinate space of input image to the atlas \cite{17}. An excellent account of segmentation approaches built on annotated atlases can be found in \cite{18}. However, creating annotations or sample pixels to define Region Of Interest would not only require expert knowledge, but are many a times infeasible. For example, tumor detection from magnetic resonance images involve investigation of more than hundred slices, that may vary significantly from patient to patient or even over time. Another growing research area that can harness the power of unsupervised segmentation is the autonomous characterization of microscopic images for identification of defect concentration in parts fabricated via advanced manufacturing methods, e.g., additively manufacturing and could be a crucial step towards realizing the materials genome initiative \cite{19}.

Towards this, several variational approaches have been reported in the literature to minimize $E(l)$ over a spatially continuous domain. Rudin, Osher and Fatemi (ROF) model \cite{20} is one of the earliest known level set-based variational approaches, followed by several variants \cite{21}, \cite{22}. However, the ROF as well as its derivative models involve minimizing an energy functional over characteristic functions of sets, resulting in a non convex model and consequently, sub-optimal convergence. Later on Chan,
Esedoglu and Nikolova presented a convex relaxation of ROF model—a continuous min-cut approach—that can be globally solved and is also the minimizer of the original ROF model after appropriate thresholding\cite{esedoglu2011convex}. Later on, authors in\cite{boykov2001fast} presented a faster and parallelizable algorithm of the continuous min-cut model using the dual capacitated max-flow model. Nonetheless, the use of expert supervision and annotated atlases remain inevitable.

Literature on unsupervised segmentation has been mostly limited. Some of the most notable works include that of normalized cuts based on spectral graph partitioning\cite{shi2000normalized}, with by several extensions thereafter\cite{shi2006normalized,boykov2001interactive,boykov2001fast}, mean shift clustering using the neighborhood pixel information\cite{comaniciu2002mean}, k-means; expectation maximization based Markov random field-based methods\cite{baron2004max} as well as mixture models\cite{zabih2005efficient,khan2008supervised}. Additionally, some of the recent works have also focused on semi-supervised methods,\emph{e.g.}, GMM guided graph cuts with bounding box as user input instead of pixel annotation\cite{schmitt2007unsupervised}, scribbles to partially annotate part of an image\cite{batra2007segmenting} and image tags that only specifies what classes are present in the image\cite{goodfellow2007boosting}. However, significant limitations exist, specially for the case of purely unsupervised approaches in terms of computational complexity, parameter selection (\emph{e.g.}, kernel function and window size in mean-shift), sensitivity to noise and over/under segmentation\cite{evers2005unsupervised}.

In this paper, we present a unsupervised approach to continuous capacitated max-flow segmentation by using an iterative maximum likelihood estimation (MLE) of the flow capacities. More specifically, we show that the maximum a posteriori (MAP) estimate of segmentation problem is analogous to the max-flow formulation for graph cuts, given the flow capacities are known. This allows for the MAP estimation of the segmentation problem. The flow capacities are then iteratively estimated by using MLE until convergence. We also present theoretical results to establish the consistency of the MAP-MLE approach, and show that the method is computationally tractable. To establish the numerical performance of the methodology, we present comparative results on benchmark tumor datasets as well as a case study on the segmentation of defect concentration in additively manufactured components fabricated with different process parameters. The paper is organized as follows:

In Section 2, we present the proposed continuous max-flow based segmentation approach. In Section 3, we show that the MAP estimate of the segmentation problem given the flow capacities is analogous to the max-flow formulation and iteratively derive the ML estimation of the flow capacities. We show that the proposed ML estimation of the flow capacity is consistent in Section 4. In section 5, we present the comparative results from experimental investigations with application in the brain tumor and defects segmentation in additively manufactured surface. We conclude the paper in Section 6.

2 \textbf{CONTINUOUS MAX-FLOW FORMULATION}

In the following we first present the continuous max-flow formulation with flow capacities for image segmentation. Subsequently, we show that the current implementations of the max-flow problem are limited and result in sub-optimal solutions due the fact that flow capacities are generally unknown and need to be estimated.

Let $\omega$ denote the continuous image domain (\emph{e.g.}, see Figure 1) and $x$ denote some pixel location in $\omega$. The problem of partitioning the continuous image domain $\omega$ into $n$ disjoint subdomains $\{\omega_i\}_{i=1}^n$ satisfying $\omega_i \cap \omega_j = \emptyset$ can be expressed as the following energy minimization problem:

\[
\min_{\{\omega_i\}_{i=1}^n} \sum_{i=1}^n \int_{\omega_i} \rho(l_i, x) dx + \alpha \sum_{i=1}^n |\partial \omega_i| \tag{2}
\]

subject to $\bigcup_{i=1}^n \omega_i = \omega; \omega_i \cap \omega_j = \emptyset, i \neq j; \alpha > 0$

where $l_i$ is the label assigned to a location $x \in \omega$ and $|\partial \omega_i|$ is the perimeter of the subdomain $\omega_i$. With reference to Figure 1(a), the subdomains $\omega_i$ may represent the defects on the surface, \emph{e.g.}, porosity, or the image background itself. The first term in Equation (2) evaluates the likelihood of assigning labels $l_i$ to locations $x \in \omega$. This is the $E_{\text{data}}(l_i)$ term in Equation (1). A classical example is the Mumford-Shah functional\cite{mumford1989optimal} that uses an $L_2$ norm to evaluate the goodness of fit of the labels, \emph{i.e.}, $\rho(l_i, x) = (l_i - f(x))^2$, where $f(x)$ is the observed label of pixel $x$. The second term $\alpha \sum_{i=1}^n |\partial \omega_i|$ enforces smoothness of the segmentation by minimizing the perimeter of each of the subdomains and is analogous to $E_{\text{smooth}}(l_i)$ in Equation (1). By imposing the smoothness penalty, we avoid over segmentation of the image. Markov random field (MRF) priors have been widely used to capture spatial smoothness. Essentially, it states that the probability of any label configuration $l = \{l_1, l_2, \ldots, l_N\}$ is given as,

\[
P(l) \propto \exp(-\sum_{x \in \omega} \sum_{y \in N(x)} V_{x,y}(f(x), f(y))) \tag{3}
\]

where $V_{x,y}(f(x), f(y))$ is the clique potential for a pair of neighboring locations $x$ and $y$. Boykov et al.,\cite{boykov2001interactive} proposed a generalized Potts model to define the clique potentials involving a pair of neighboring locations as,

\[
V_{x,y}(x, y | N(x)) = c \exp \left(-\frac{\sum_{y \in N(x)} |f(x) - f(y)|^\sigma}{\sigma} \right) \tag{4}
\]
Let $u_i(x), i = 1, \ldots, n$ denote an indicator function such that,
$$
u_i(x) = \begin{cases} 1 & x \in \omega_i \\ 0 & x \notin \omega_i \end{cases} \quad (5)$$
we have, $|\partial \omega_i| = \int_\omega |\nabla u_i(x)|dx \forall i = 1, \ldots, n$. Therefore, we can rewrite the energy functional as,
$$\min_{\{u_i(x)\} \in \{0,1\}} \sum_{i=1}^n \int_\omega u_i(x)\rho(l_i,x)dx + \alpha \sum_{i=1}^n \int_\omega |\nabla u_i(x)|dx \quad (6)$$
subject to $\sum_{i=1}^n u_i(x) = 1 \forall x \in \omega, \alpha > 0$.

Note that the formulation in Equation (6) is non-convex due to the binary configuration of $u_i(x)$. Several convex approximations of Equation (6) has been proposed in the literature by relaxing $u_i(x) = [0,1]$ to the interval $[0,1]$. Authors in [9] showed that the optimal solution to Equation (6) can be obtained by thresholding the resulting convex problem. However, the numerical algorithms for the model in Equation (6) suffer from the non-smoothness of the total variation term $\sum_{i=1}^n \int_\omega |\nabla u_i(x)|dx$. To overcome this, we show in the following, that the dual of the energy minimization problem in Equation (6) is analogous to the continuous max-flow problem studied in [23], [33]. Towards this, we consider a specific formulation of Equation (6) as derived in [9],

$$\min_{\{u\} \in [0,1]} \int_\omega (1-u)C_sdx + \int_\omega uC_tdx + \alpha \int_\omega |\nabla u|dx \quad (7)$$

Note the relaxation on $u_i(x)$ to $[0,1]$. Using the fact that $\alpha \int_\omega |\nabla u|dx = \max_{|p(x)| \leq C(x)} \int_\omega \text{div}(p)|dx$ (see [37]), we rewrite Equation (7) as the following min-max problem,

$$\min_{\{u\} \in [0,1]} \int_\omega (1-u)C_sdx + \int_\omega uC_tdx + \alpha \int_\omega |\nabla u|dx$$

$$= \max_{\{u\} \in [0,1]} \max_{p(x) \in [0,C_s]} \int_\omega (1-u)p_s + up_t + \text{div}(p)|dx$$

$$= \max_{p(x) \leq C_s(x)} \min_{|p| \leq C(x)} \int_\omega p_s + u(\text{div}(p) - p_s + p_t)|dx (\text{subject to } p_s(x) \leq C_s(x), p_t(x) \leq C_t(x), |p(x)| \leq C(x)) \quad (12)$$

Now, we can write the dual of Equation (6) as,

$$\max \int_\omega p_sdx \quad \text{subject to } p_s(x) \leq C_s(x)$$

$$p_t(x) \leq C_t(x)$$

$$|p(x)| \leq C(x)$$

$$\text{div}(p) - p_s + p_t = 0$$

The resulting formulation in Equation (8) is the continuous analogue of the max-flow problem (also see [12] for a more formal derivation). To understand this max-flow formulation in the continuous domain, we refer to Fig. [6]. In the continuous image plane $\omega$, each pixel $x$ is connected to the source $s$ and sink $t$. Following [12], we consider that each pixel $x \in \omega$ is associated with three different flows: the source flow $p_s(x) \in \mathbb{R}$, the sink flow $p_t(x) \in \mathbb{R}$ and the spatial flows $p(x) \in \mathbb{R}^2$. Here, source and sink flow fields are directed from the source $s$ to each point $x \in \omega$ and from each point $x$ to the sink $t$, respectively. The spatial flow field is characterized by the undirected flow through $x$—capturing the strength of interaction with the neighborhood locations. Each of the flow fields are constrained by their respective capacities as,

$$p_s(x) \leq C_s(x), p_t(x) \leq C_t(x), |p(x)| \leq C(x) \quad (9)$$

In addition, the flow fields satisfy the flow conservation constraint, i.e.,

$$\text{div}(p) - p_s + p_t = 0 \quad (10)$$

Thus the continuous max-flow problem can be defined as the maximization of the total flow from the source $s$ as,

$$\max \int_\omega p_sdx \quad \text{subject to the constraints in Equations (9)-(10).}$$

To solve the minimization problem in Equation (8), we first write the corresponding Lagrangian using the multiplier $\lambda$ as,

$$\max_{\lambda} \int_\omega p_sdx + \int_\omega \lambda(\text{div}(p) - p_s + p_t)|dx \quad (12)$$

subject to $p_s(x) \leq C_s(x), p_t(x) \leq C_t(x), |p(x)| \leq C(x)$

With this, we can write the following result,

**Proposition 1** Optimizing the dual variables $p_s(x), p_t(x), p(x)$ in Equation (12) is equivalent to minimizing the corresponding primal problem in Equation (6).

Proof of the proposition follows from the Equations (7)-(8). Now we can solve the maximization problem in Equation (12) by using the augmented Lagrangian approach presented in [23] as,

$$L(p_s,p_t,p,\lambda) = \int_\omega p_sdx + \int_\omega \lambda(\text{div}(p) - p_s + p_t)|dx$$

$$- \frac{c}{2}||\text{div}(p) - p_s + p_t||^2 \quad (13)$$

where $c > 0$. See Algorithm 1 [23] for solving the augmented Lagrangian $L(p_s,p_t,p,\lambda)$ using the projection-gradient de-
As mentioned earlier, in unsupervised max-flow problems, the flow capacities corresponding to the source, sink as well as the spatial field flows are set to a fixed value a priori by the user or requires user input to define foreground and background [ref]. Nonetheless, the capacities are generally unknown and may require trial and error or prior training. Additionally, for images with low SNR, fixed capacity values for all the pixels \( x \in \omega \) would cause the solution to trap in the local minima, thus failing to determine the global optimum (see Figure [3] for an example). To overcome these challenges, we present an approach to define spatially varying flow capacities. In the following, we first formulate the MAP estimate of the segmentation problem, i.e., \( \lambda \) for a given estimate of the flow capacities and show that the MAP formulation is analogous to solving the continuous max-flow problem. With current estimate of \( \lambda \), we subsequently update the flow capacities in the next iteration using a maximum likelihood approach.

**Algorithm 1: Continuous max-flow using gradient projection**

1. Initialize: \( p^0_s(x), p^0_t(x), p^0_l(x), \lambda^0 \) and \( k = 0 \)
2. Repeat
   3. For every subdomain \( \omega_i, i = 1, 2, \ldots, n \) do
      4. Update the spatial flows
         \[
         p^{k+1}_i \leftarrow p^k_i - \gamma \nabla (\text{div}(p^k_s - D^k_i(x))),
         \]
         where \( D^k_i(x) \leftarrow (p^k_s + \lambda^k_i - 1) \)
      5. Update the sink flows;
      6. \( F^k_i(x) \leftarrow (p^k_s + \lambda^k_i - 1) \)
      7. Update the source flows;
      8. \( G^k_i \leftarrow (p^{k+1}_s + \lambda^k_i - 1)(x)/c \), where \( p^{k+1}_s \leftarrow (1 + c \sum_{i=1}^{n} C_i^k(x))/nc \)
      9. Update the multipliers;
     10. \( \lambda^{k+1}_i \leftarrow \lambda^k_i - c(\text{div}(p^{k+1}_s + (p^{k+1}_s + p^{k+1}_t)) \end{algorithm

### 3 Maximum a Posteriori/Maximum Likelihood Estimation

As mentioned earlier, in unsupervised max-flow/min-cut problems, the value of capacities corresponding to the source, sink as well as the spatial field flows are set to a fixed value a priori by the user, introducing subjectivity into the segmentation. Assuming a constant capacity constraint ignores the neighborhood information, and therefore the solution is highly sensitive to the noise. In fact, convergence of the label configuration \( \lambda \) in most max-flow/min-cut based approaches is dependent on the source, sink and spatial flow field capacities. We show this in the following result.

**Theorem 1** Let \( \lambda \) be some label configuration (but not the optimal) with \( p_s(x), p_t(x) \) and \( p(x) \) as the source, sink and spatial flows, respectively. Let the corresponding source, sink and spatial capacities be equal to the optimal flows, i.e., \( C_s(x) = p^*_s(x), C_t(x) = p^*_t(x) \) and \( C(x) = p(x) \). Then the label configuration \( \lambda \rightarrow \lambda^* \), where \( \lambda^* \) is the true configuration.

**Proof 1** To derive the proof, we first consider the following subproblems:

\[
\begin{align*}
    f_s(x) & = \sup_{p_s(x) \leq C_s(x)} (1 - \lambda(x))p_s(x) \\
    f_t(x) & = \sup_{p_t(x) \leq C_t(x)} \lambda(x)p_t(x)
\end{align*}
\]

Let \( p^*_s(x) \) be the optimal solution of (14). Under optimality, let \( f_s(x) < \lambda^s(x) \). Similarly decreasing the source capacity decreases the source flow field, i.e., \( p^*_s(x) \) and therefore, \( \lambda^s < \lambda^* \). In either case, the optimality is lost. In a similar way, we can show the case for sink and spatial flows.

To overcome these challenges, we estimate spatially varying flow capacities. We first formulate the MAP estimate of the segmentation problem, i.e., \( \lambda \) and show that the estimation of the MAP is analogous to solving the continuous max-flow problem.

For each of the image subdomains \( \{\omega_i\}_{i=1}^{n} \subset \omega, \omega_i \cap \omega_j = \emptyset \), we assign the source, sink and spatial capacities to each of the subdomains \( \omega_i \) as \( C_s(x \in \omega_i), C_t(x \in \omega_i) \) and \( C(x \in \omega_i) \). We consider the capacities to be spatially varying over the domain \( \omega \), but are fixed within each of the subdomains. Given the continuous image domain, \( \omega \) and the flow capacities, the objective is to find the label configuration \( \lambda \) that maximizes the posterior probability given as,

\[
\hat{\lambda} = \arg \max_{\lambda} P(\lambda | C_s(x), C_t(x), C(x), \omega)
\]

As we mentioned earlier, the flow capacities are not known a priori and is difficult to estimate, especially for images with low SNR ratio (e.g., see Figure [3]). In light of this, we use an iterative learning approach to estimate the flow capacities, \( C = \{ C_s(x), C_t(x), C(x) \} \) and \( \hat{\lambda} \) as,

\[
\begin{align*}
    \lambda^{(t)+1} & = \arg \max_{\lambda} P(\lambda | C(x), \omega) \\
    C^{(t)+1} & = \arg \max_{C} P(\lambda^{(t)+1} | C, \omega)
\end{align*}
\]

Given the pixel intensities in the image domain \( \omega \), the joint distribution of the flow capacities \( C \) and the label
configuration $\lambda$ can be written using the Bayes theorem as,

$$P(C, \lambda | \omega) \propto P(\omega | C, \lambda) P(\lambda | C) P(C)$$  \hspace{1cm} (19)

Here, the first term $P(\omega | C, \lambda)$ is nothing but the likelihood of the labels $\lambda$ and the capacities $C$ given the image domain $\omega$, $P(\lambda | C)$ is some prior over the image domain given the capacities $C$ and $P(C)$ denotes the prior distribution representing the capacities. To appropriately define the prior distribution over the capacities it is important to understand that the flow capacities should have similar characteristics as that of the optimal flow. That is, the capacities should be smoothly varying over the image domain as well as capture the spatial correlation between the image pixels. In fact, for an optimal solution, we note from Theorem 1 that $C^*_s(x) = p^*_s(x), C^*_t(x) = p^*_t(x)$ and $C^*_\omega(x) = p^*(x)$. In the following, we focus only on the estimation of the source capacity $C_s(x)$. Similar results may be obtained for the sink and spatial flow field capacities.

With that, one of most commonly used prior to incorporate the spatial smoothness is the Markov random field (MRF) prior given as,

$$p(C_s(x)) = \frac{1}{Z} \exp{-\frac{1}{T} \sum x \in \omega \sum m \in N(x) (C_s(x|\omega_i) - C_s(m))^2}$$  \hspace{1cm} (20)

where $Z$ is a normalizing constant, $T$ is a scale factor and $U(C_s(x))$ is a smoothing function and controls the spatial correlation between the pixels in a given neighborhood. The prior follows from the Hammersley-Clifford Theorem [ref] when assuming local Markovian property over the flow field, that is, $p(C_s(x)|C_s(\omega \sim x) = p(C_s(x|C_s(x \in N(x))))$. It may be noted that local Markovian property is quite natural for random field defined over an image domain.

Several choices of $U(C_s(x))$ has been proposed in the literature [38, 39]. A typical choice for $U(C_s(x))$ is the neighborhood based Gauss-Markov random field prior where $U(C_s(x))$ is given as:

$$U(C_s) = \sum x \in \omega \sum m \in N(x) (C_s(x|\omega_i) - C_s(m))^2$$  \hspace{1cm} (21)

Authors in [40] proposed the following form for $U(C_s(x))$ to penalize the larger values of $x$ while increasing the robustness to the outliers,

$$U(C_s(x)) = \sum x \in \omega \sum m \in N(x) \left(1 + \frac{1}{(C_s(x) - C_s(m))^2}\right)^{-1}$$  \hspace{1cm} (22)

However, as pointed out it [1], the functional form of $U(C_s(x))$ in Equations (21)&(22) result in complex log-likelihood functions making it computationally intensive. To overcome the computational complexity, we use the prior as proposed in [1].

$$U(C_s(x)) = -\sum x \in \omega \sum m \in N(x) G^{(t)}(x \in \omega_i) \log \left(C^{(t+1)}_s(x \in \omega_i)\right)$$  \hspace{1cm} (23)

where $G^{(t)}(x \in \omega_i)$ is defined as

$$\exp{\left(\frac{\beta}{2|N(x)|} \sum m \in N(x) \left(p_s^{(t)}(m \in \omega_i) + C_s^{(t)}(m \in \omega_i)\right)\right)}$$  \hspace{1cm} (24)

where $N(x)$ is the cardinality of the neighborhood of $x$, $p_s^{(t)}$ and $C_s^{(t)}$ are the current estimates of the source flow field and source flow capacity, and $\beta$ is a smoothing constant. An inherent advantage of using the smoothing function $U(C_s(x))$ given in Equation (23) is that when maximizing the log-likelihood, the derivative is dependent only on the term $C_s(x)$; contributing to the computational efficiency. In the present implementations, we set the value of $\beta = 5$ and the neighborhood size as $5 \times 5$ over $\omega$ such that $|N(x)| = 25$. Different values result in different convergence rates.

Returning to Equation (19), we first write the corresponding log-likelihood function as,

$$L(C|\lambda, \omega) = \log(P(C, \lambda | \omega))$$  \hspace{1cm} (25)

$$= \log(P(\omega | C, \lambda)) + \log(P(\lambda | C)) + \log(P(C))$$  \hspace{1cm} (26)

$$= \log(P(\omega | C, \lambda)) + \log(P(\lambda | C)) - \log(Z) - U(C_s) + \frac{T}{10}$$

Here, the first term represents the likelihood of the capacity given the label configuration $\lambda$ and the image domain $\omega$. In other words, it imposes a penalty for every incorrect assignment of the labels. Assuming each of the observations are independently and identically distributed we can write,

$$P(\omega | C, \lambda) \propto \exp{-D(C_s(x), \lambda, \omega)}$$  \hspace{1cm} (27)
such that, 
\[ D(C_s(x), \lambda, \omega) = \sum_{i=1}^{n} \int_{\omega} u_i(x) \rho(l_i, x) dx \]
where \( \rho(l_i, x) \) is some loss function for assigning label \( l_i \) to location \( x \) (see (2)). The second term in Equation (25) represents the prior over the label configuration \( \lambda \). We consider a gradient based smoothness prior for \( P(\lambda | \omega) \) given as,
\[ P(\lambda | \omega) \propto \exp \left( \sum_{x \in \omega} \sum_{m \in N(x)} -V_{x,m} \right) \]  
(28)
where \( V_{x,m} \) is the smoothness penalty and is given as
\[ V_{x,m} = \alpha \sum_{i=1}^{n} \int_{\omega} |\nabla u_i(x)| dx \]
Note that we only provide a generic form for \( D(C_s(x), \lambda, \omega) \) and \( V_{x,m} \). More specialized forms can also be used. Using, the generalized penalty function and the gradient based regularization, as introduced in Equation (4), we rewrite Equation (25) as,
\[
\log \left( \exp \left( \sum_{i=1}^{n} \int_{\omega} u_i(x) \rho(l_i, x) dx \right) \right) \\
+ \log \left( \exp \left( \alpha \sum_{i=1}^{n} \int_{\omega} |\nabla u_i(x)| dx \right) \right) - \log(Z) - \frac{U(C_s)}{T} \\
= \log \left( \prod_{i=1}^{n} \exp \left( \int_{\omega} u_i(x) \rho(l_i, x) dx \right) \right) + \\
\log \left( \prod_{i=1}^{n} \exp \left( \alpha \int_{\omega} |\nabla u_i(x)| dx \right) \right) - \log(Z) - \frac{U(C_s)}{T} \\
= \left( \sum_{i=1}^{n} \int_{\omega} u_i(x) \rho(l_i, x) dx + \alpha \sum_{i=1}^{n} \int_{\omega} |\nabla u_i(x)| dx \right) \\
- \log(Z) - \frac{U(C_s)}{T} \tag{30}
\]
Note that maximizing the log-likelihood function in Equation (29) amounts to solving the max-flow problem as in Equation (6) given the capacities are known. We use the Algorithm 1 as proposed in [23] to find the optimal \( \lambda, p_s(x), p_t(x) \) and \( p(x) \). With the current estimate of \( \lambda^{(t)}, p_s^{(t)}(x), p_t^{(t)}(x) \) and \( p^{(t)}(x) \), we solve for the optimal value of \( C_s(x) \) by maximizing \( L(C_s(x)|\lambda, \omega) \) with respect to \( C_s(x) \), i.e., we solve the following problem,
\[
\frac{\partial L(C_s(x)|\lambda, \omega)}{\partial C_s(x)} = 0 \tag{31}
\]
We solve for the next step \( C_s^{(t+1)}(x \in \omega_i) \) and is given as,
\[
C_s^{(t+1)}(x \in \omega_i) = G^{(t)}(x \in \omega_i) + \left( 1 - p_s^{(t)}(x \in \omega_i) \right) \tag{32}
\]
The algorithm for simultaneous update of \( p_s^{(t)}(x) \) and \( C_s^{(t)}(x) \) is given in Algorithm 2. Other flow capacities can be updated in a similar way.

4 CONVERGENCE AND CONSISTENCY

Algorithm 2: Iterative MLE of source, sink and spatial capacities

 Initialize: Initialize the parameters \( \lambda^{(0)}, p_s^{(0)}(x), p_t^{(0)}(x), p^{(0)}(x) \) and \( C^{(0)}(x) \)

 repeat

 1. Solve the continuous max-flow in Equation (12) to estimate \( \lambda^{(t)}, p_s^{(t)}(x), p_t^{(t)}(x) \) and \( p^{(t)}(x) \);

 2. Evaluate the current estimate of \( G^{(t)}(x \in \omega_i) \) in Equation (24);

 3. Update the capacities as

\[
C_s^{(t+1)}(x \in \omega_i) = G^{(t)}(x \in \omega_i) + \left( 1 - p_s^{(t)}(x \in \omega_i) \right)
\]

\[
C_t^{(t+1)}(x \in \omega_i) = G^{(t)}(x \in \omega_i) + p_t^{(t)}(x \in \omega_i)
\]

\[
C^{(t+1)}(x \in \omega_i) = G^{(t)}(x \in \omega_i) + p^{(t)}(x \in \omega_i)
\]

 until convergence;

To check that the algorithm converges, we analyze the likelihood function given in Equation (29) after every iteration of MAP and ML estimation. Let the likelihood function after some iteration \( t \) be \( L^{(t)} \). In the MAP estimate of the segmentation, we solve the max-flow problem, therefore, the log-likelihood at \( (t+1) \) iteration will increase (or remain same), i.e., \( L^{(t+1)} \geq L^{(t)} \) (this is evident with \( Z = 1 \) and \( U(C_s) \) as given in Equation (25)). The next step involves updating the capacities using the ML estimates. In this step the likelihood function further decreases or remains same as this step only changes the flow capacities without changing the estimated labels. Therefore, every instance of MAP and ML estimations result in the increase in the likelihood function. To show that the algorithm consistently estimates the flow capacities, we refer to the classical consistency results of MRF prior given in [41]. In this regard, we first consider the following lemma.

Lemma 1 The function \( L(C_s(x)|\lambda, \omega) \) is identifiable and is a concave objection function.

Proof 2 Identifiability: We show this for the case of \( C_s(x) \). From Equations (7,23&24), we have
\[
L(C_s^{(t+1)}(x)|\lambda, \omega) = k + \sum_{x \in \omega} \sum_{m=1}^{n} G^{(t)} \log \left( C_s^{(t+1)}(x) \right)
\]
where \( k \) is some constant. Clearly \( L(C_s^{(t+1)}(x)|\lambda, \omega) \) is one-to-one and therefore, identifiable.

Concavity: As logarithm is concave in the domain of \((0,1)\), the linear transform \( k + \sum_{x \in \omega} \sum_{m=1}^{n} G^{(t)} \log(C_s^{(t+1)}(x)) \) is also concave.

With the conditions stated in Lemma 1, we state the following consistency Theorem as derived in [41].

Theorem 2 Given \( L(C_s(x)|\lambda, \omega) \) is identifiable and concave. The MLE \( L(C_s(x)) \) is always consistent, i.e.,
\[
P(\|L(C_s(x)) - L(C_s(x))\| \geq \epsilon) \leq c \exp(-\delta)
\]
See [41] for the proof. Based on aforementioned results, we note that every instance of MAP and ML estimations result in the increase in the likelihood function, ensuring that
the algorithm consistently converges. In the next section, we present the implementation results and comparison with some state-of-the-art methods.

5 EXPERIMENTAL RESULTS

5.1 Case studies and evaluation metrics

To test the efficacy of the proposed methodology, we investigate two real-world case studies, including Brain Tumor Image Segmentation (BRATS, datasets from 2013 and 2015) and defect localization in additively manufactured (AM) surfaces gleaned from electron microscopy. An important distinction between the two datasets is that the former is sufficiently large to enable supervised learning. In contrast, high resolution electron microscopic images are extremely costly to capture from AM surface mostly due to the uncertainties in the surface morphology, limiting the availability of such images. As most of the segmentation approaches in the past are either supervised or involves user interaction and manual parameter setting, we compare the segmentation results from each of the cases studies with different state of the art algorithms.

For the first case study, we refer to the algorithms reported in the Medical Image Computing and Computer Assisted Intervention Society (MICCAI) proceedings 2013 and 2015 respectively [42]. In the 2013 MICCAI proceedings, 20 algorithms were reported that included (5) generative, (13) discriminative and (2) generative-discriminative algorithms. In the 2015 BRATS dataset, (3) generative, (2) discriminative and (7) neural network/deep learning approaches were reported. Additionally, we also implement the state of the art unsupervised algorithms, including k-means, expectation maximization (blobworld) [43], Gaussian mixture model (GMM), GMM with spatial regularization (SCGMM) [1], mean shift [27], normalized cuts [2], and hierarchical image segmentation [26]. Due to the limited dataset in the second case study, we compare the performance with only unsupervised algorithms. Among the methods tested, the ones that did not perform or resulted in poor segmentation are not reported.

To quantitatively compare the performance of the segmentation results, we refer to the standard Dice score and the Hausdorff distance. The Dice score is given as

$$\text{Dice} = 2 \frac{|\hat{R} \cap R|}{|\hat{R}| + |R|}$$

(33)

where $\hat{R}$ and $R$ represents the estimated lesion region and the expert segmentation, respectively and $|.|$ represents the size of the domain. Dice score measures the areal overlap or agreement between the segmented and ground truth. The Hausdorff distance given in Equation (34) instead measures the surface distance between the segmentations,

$$\text{Haus}(\hat{R}, R) = \max \left\{ \sup_{i \in \hat{R}} d(i, R), \sup_{j \in R} d(j, \hat{R}) \right\}$$

(34)

where $d(j, \hat{R})$ is the shortest Euclidean distance between $j$ and $\hat{R}$. Maximum over the supremum distances make Hausdorff distance highly sensitive to the outliers in detection. To control this, we use the 95th percentile of the Hausdorff distance as suggested in [42].

5.2 Brain Tumor Segmentation

The World health Organization has classified brain and central nervous system tumors into more than 120 categories, along with a grading scale specifying the degree of malignancy (on the scale of I-IV). However, almost 80% of the malignant tumors fall in the gliomas category alone. Despite significant advances over the past few years, the diagnosis of the gliomas remain limited. Neuroimaging offers a noninvasive approach to evaluate the progression of the tumorous tissues, thus allowing timely intervention and controlled treatment. However, significant challenges arise when diagnosing the MR images as the lesions (tumorous regions) are defined only in terms of the contrast changes with respect to the surrounding normal tissues. This often causes significant variations among the expert segmentation mainly because of the intensity gradient from the tumorous to non-tumorous regions. In addition, these lesions may vary significantly in terms of size, shape and localization.
from patient to patient, as well as across subsequent stages of tumor growth, requiring large training datasets for an effective supervised learning approach.

In light of this, we study the segmentation of the brain tumor from fluid-attenuated inversion recovery (FLAIR) MR scans that highlights the differences in water relaxation properties of the tissues, and is effective in enhancing the peritumoral edema—a characteristic feature of malignant gliomas. The FLAIR images used in this study are acquired from two different instances of the BRATS datasets, BRATS 2013, consisting of a total of 20 high grade (HG) and 10 low grade (LG) glioma cases, and BRATS 2015, consisting of 274 HG and 54 LG gliomas. Each of the datasets contain annotated ground truth delineated by clinical experts.

Figure (5) shows a representative segmentation for HG and LG gliomas along with the comparative results from other unsupervised algorithms. Although there is some mismatch between the segmented and the ground truth regions for the presented algorithm, it is able to outperform the other unsupervised approaches. The overall performance of the presented algorithm is 72% and 76% on the Dice score for the 2013 and 2015 datasets, respectively. The 95% Hausdorff measure measured for the two separate instances are 16.08 and 14.05. The Dice score and the Hausdorff measure of all the methods including the supervised (automated as well as manual initialization) as well as unsupervised algorithms tested on the 2013 dataset are summarized in Figures 6(a&b). Clearly, our unsupervised approach outperforms most of the supervised algorithms, specifically in terms of the average values of Dice score and Hausdorff measure (black squares inside the box plot indicate the average).

For BRATS 2015 competition no test result was available at the time of writing, so we report only the training results of the algorithms reported in the MICCAI 2015 proceedings. This is shown in Table 1. Indeed, most of the deep learning based method outperform the performance of the our algorithm. However, it must be noted that these are the Dice scores on training dataset.

### Table 1: Dice score comparison for the BRATS 2015 dataset

| Approach                        | Dice score     | Author     |
|---------------------------------|----------------|------------|
| Unsupervised min cut            | 75.6±10.5      | Iquebal    |
| Generative with shape prior     | 77±19          | Agn [44]   |
| Generative-Discriminative       | 88.5±20        | Bakas [44] |
| Expectation Maximization        | 68             | Haack [44] |
| Random Forests                  | 84             | Maier [44] |
| Random Forests                  | 80.7           | Malmi [44] |
| Conditional Random Fields       | 88.7±35        | Meier [44] |
| Convolutional Neural Network    | 88             | Havaei [44]|
| Convolutional Neural Network    | 81             | Dvorak [44]|
| Convolutional Neural Network    | 86             | Pereira [44]|
| Convolutional Neural Network    | 67             | Rao [44]   |
| Convolutional Neural Network    | 81.41±9.6      | Vaidhya [44]|
| Convolutional Neural Network    | 87.55±6.72     | Wang [13]  |

### 5.3 Defect concentration in additively manufactured components

In this case study, we explore a rather newly emerging field of materials and process discovery. Recent advances in the manufacturing technologies, specially additive manufacturing (AM, layer by layer deposition of metal powder to fabricate complex free-form surfaces) have revolutionized the landscape of fabricating industrial components and parts. With newer technologies, newer challenges arise. In the case of AM, although, it is capable of fabricating components with minimum time and material waste, overall functional integrity of AM components is considered much inferior to those realized with conventional manufacturing process chains, especially under real world dynamic loading conditions. Defects, such as pores, undiffused powder, geometric distortions, surface cracks and non-equilibrium microstructures significantly deteriorates the mechanical performance...
Fig. 6: Comparative results of different algorithms tested for the segmentation of defects. (a) Original image for sample A (b) k-means with 2 clusters (c) Gaussian mixture model with expectation maximization (d) spatially constrained Gaussian mixture model with k-means initialization and (e) mean shift (f) the proposed method.

and overall functional integrity of the components. Developments of in situ imaging technologies allow monitoring of manufacturing processes at every conceivable resolution of interest. However, analysis of the images recorded from as-fabricated components to characterize the type and concentration of defects is not a trivial task. These microscopic images contain significant uncertainties in the shape and morphology, e.g., an undiffused metal powder appears very similar in shape to spherical pores. In addition, the signal to noise ratio of the images is typically very low and consists of under- and overexposed regions. Figures 1(a&b) shows the morphology of defects formed during two different additive manufacturing processes. Only a few studies have been reported in the literature that utilizes in situ sensor data for the characterization of defects, e.g., see [45].

We test the performance of our approach on two surfaces that were fabricated using different process parameters in a laser based AM process called selective laser melting (SLM). The process parameters for the first surface (sample A, Figure 5(a)) was set as: laser power = 165 W, laser scanning speed = 138 mm/s and relative density = 99.5% and the second workpiece (sample B, Figure 6(a)) was: laser power = 85 W, laser scanning speed = 71 mm/s and relative density = 96.4%, respectively. All other process parameters, such as hatching distance (mm), layer thickness (mm), etc. were kept fixed.

In the current study we focus on the concentration of two different type of defects namely, balling effect and porosity. Balling effect is a complex metallurgical process that originates due to sub-optimal process parameters during the SLM process as well as the properties of the material powder. It causes the liquid scan track during SLM to break and result in the formation of spherical particles that eventually get trapped, causing inhomogeneous deposition of the powder in the next build layer. Authors in [30] showed that by increasing the laser power, the concentration of defects due to balling effect decreases. The second type of defects are the pores. Pores are the voids on surface that are formed when gas particles trapped in the melt-pool (liquefied metal during deposition) escapes. These are generally tracked by observing the dark spherical/oval shaped features on the surface. Controlling porosity as well as balling effect is critical to avoid significant costs, as these defects may reduce the fatigue strength of material by more than 4 times, resulting in early, unexpected failure [46].

Figure 6(a) and 7(a) shows the representative surface from sample A and B, respectively. Clearly sample A with higher laser power (165 W) has relatively lower defect concentration as compared to sample B with low laser power (85 W). As the data size is limited, we compare our segmentation results only with that of other unsupervised segmentation methods, including k-means, Gaussian mixture model (GMM), and spatially constrained Gaussian mixture model (SCGMM) [1] and mean-shift [27]. Note that we ignored comparisons with graph cut approaches as some of the the existing implementations either required user input to define the foreground and background or did not converge. For the first sample, segmentation results are presented in Figures 6(b-f). The image contains significant noise due to the shadow and overexposure resulting due to uneven morphology of the AM surfaces, which makes the segmentation process challenging. The results obtained from the k-means as shown in Figure 6(b) contains a lot of noise. There is, however, some improvement when segmented with the GMM as shown in Figure 6(c). Similarly with SCGMM using k-means initialization is able to reduce
TABLE 2: Dice score and Hausdorff measure of various unsupervised approaches implemented for defect segmentation

| Approach                  | Dice score | Hausdorff Measure |
|---------------------------|------------|-------------------|
| Unsupervised min cut      | 70.56      | 32                |
| mean shift                | 58.6       | 245               |
| Gaussian mixture model     | 22.16      | 287               |
| Spatially constrained GMM  | 2.6        | 219               |
| k-means                   | 1.4        | 321               |

the noise (see the white region, in Figure 6(d)). The segmentation from the mean shift algorithm (Figure 6(e)) is able to detect most of the pores but fails to detect the defect induced by the balling effect. The segmentation obtained with the proposed method is clearly able to identify all the regions of pores as well as the bailing effect (Figure 6(f)).

In the second sample (B), we note that the noise is significantly higher as compared to sample A, mainly due to the rougher surface morphology. Three instances of bailing effect were identified manually as shown by the arrows, along with multiple pores. Clearly, the segmentation results obtained from k-means (Figure 6(b)), GMM (Figure 6(c)) and SCGMM (Figure 6(d)) as well as mean shift (Figure 6(e)) all capture mostly the noise in the images. The segmentation obtained using the proposed method is clearly able to remove all noise and is selectively able to capture the defects (Figure 6(f)). However, for the case of bailing effect, the algorithm is able to identify only two bailing effects clearly. Nonetheless, the segmentation is a significant improvement over the standard state of the art methods.

Estimating defect concentration would enable understanding the effect of different process parameters on the build quality, and therefore determine the optimal parameter settings. We verified the experimental observations reported by [30], that increase in the laser power results in less defect concentration. We validated this observation by estimating the proportion of defect area. The defect concentration on sample A is \( \sim 1\% \) and sample B is \( \sim 4.39\% \) as estimated from the segmented images, that is in accordance with the observations.

6 CONCLUSIONS

In the present work, we developed an graph coloring approach to segment noisy images using a capacitated continuous max-flow approach augmented with a maximum likelihood estimation of the flow capacities. We show that the maximum \textit{a posteriori} estimate of the segmentation problem is analogous to the max-flow formulation given the flow capacities are known. The flow capacities are then estimated using the maximum likelihood estimation using a Markov random field prior over the neighborhood structure. The comparative results in an industrial setting established the performance of the present approach, as the method was able to selectively segment the defects ignoring the noise appearing due to shadow and presence of over/under exposed regions. Current and future works are focused on establishing the theoretical guarantees on the convergence rates and more informative priors.

ACKNOWLEDGMENTS

REFERENCES

[1] T. M. Nguyen and Q. J. Wu, “Fast and robust spatially constrained gaussian mixture model for image segmentation,” \textit{IEEE transactions on circuits and systems for video technology}, vol. 23, no. 4, pp. 621–635, 2013.
[2] J. Shi and J. Malik, “Normalized cuts and image segmentation,” *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, no. 8, pp. 888–905, 2000.

[3] R. Zabih and V. Kolmogorov, “Spatially coherent clustering using graph cuts,” in *Computer Vision and Pattern Recognition*, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on, vol. 2. IEEE, 2004, pp. II–II.

[4] F. Forbes and N. Peyrard, “Hidden markov random field model selection criteria based on mean field-like approximations,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, pp. 1089–1101, 2003.

[5] H. Greenspan, G. Dvir, and Y. Rubner, “Context-dependent segmentation and matching in image databases,” *Computer Vision and Image Understanding*, vol. 93, no. 1, pp. 86–109, 2004.

[6] J.-S. Kim and K.-S. Hong, “Color-texture segmentation using unsupervised graph cuts,” *Pattern Recognition*, vol. 42, no. 5, pp. 735–750, 2009.

[7] Y. Boykov and M.-P. Jolly, “Interactive graph cuts for optimal boundary & region segmentation of objects in nd images,” in *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, vol. 1. IEEE, 2001, pp. 105–112.

[8] Y. Boykov, O. Veksler, and R. Zabih, “Fast approximate energy minimization via graph cuts,” *IEEE Transactions on pattern analysis and machine intelligence*, vol. 23, no. 11, pp. 1222–1239, 2001.

[9] T. Chan, S. Esedoglu, and M. Nikolova, “Algorithms for finding global minimizers of image segmentation and denoising models,” *SIAM journal on applied mathematics*, vol. 66, no. 5, pp. 1632–1648, 2006.

[10] S. Geman and D. Geman, “Stochastic relaxation, gibbs distributions, and the bayesian restoration of images,” *IEEE Transactions on pattern analysis and machine intelligence*, no. 6, pp. 721–741, 1984.

[11] M. Klotz, T. Schoenemann, K. Kolev, M. Schikora, and D. Cremers, “An experimental comparison of discrete and continuous shape optimization methods,” in *European Conference on Computer Vision*. Springer, 2008, pp. 332–345.

[12] G. Strang, “Maximal flow through a domain,” *Mathematical Programming*, vol. 26, no. 2, pp. 123–143, 1983.

[13] D. Cohen-Or, E. Grinszpan, C. Liang, J. Snyder, and D. Breen, “Interactive boundary extraction using iterated graph cuts,” in *ACM transactions on graphics (TOG)*, vol. 23, no. 3. ACM, 2004, pp. 309–314.

[14] J. Xu, A. G. Schwing, and R. Urtasun, “Tell me what you see and i will show you where it is,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 3190–3197.

[15] T. Kohli, P. Kohli, and B. Triggs, “Multi-label image segmentation,” in *Computer Vision–ECCV 2010*. Springer, 2010, pp. 222–235.

[16] A. Shekhovtsov, A. Shekhovtsova, and K. S. Trivedi, “Multi-label image segmentation using sequential inference,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 7, pp. 1298–1312, 2016.

[17] J. Balachandran, L. Jin, J. S. Anchell, C. A. Bridges, and P. Ganesh, “Defect genome of cubic perovskites for fuel cell applications,” *The Journal of Physical Chemistry C*, vol. 121, no. 48, pp. 26 637–26 647, 2015.
high resolution imaging." Additive Manufacturing, vol. 21, pp. 517–528, 2018.

[46] S. Tammas-Williams, P. Withers, I. Todd, and P. Prangnell, “The influence of porosity on fatigue crack initiation in additively manufactured titanium components,” Scientific reports, vol. 7, no. 1, p. 7308, 2017.