

## 融合时序信息的激光焊接熔透状态识别方法

刘天元<sup>1</sup>, 鲍劲松<sup>1\*</sup>, 汪俊亮<sup>1</sup>, 顾俊<sup>2</sup>

<sup>1</sup> 东华大学机械工程学院智能制造研究所, 上海 201600;

<sup>2</sup> 上海市激光技术研究所, 上海 200235

**摘要** 熔池/匙孔区域包含丰富的激光焊接质量信息, 基于该区域对熔透状态进行准确识别对于激光焊接的在线控制至关重要。针对熔池区域存在蒸汽和飞溅等干扰且相邻的熔透状态之间特征区分度低带来的模式识别困难的问题, 提出了一种融合时序信息的激光焊接熔透状态识别方法。所提方法以熔池/匙孔图像序列作为底层输入, 首先通过两个卷积层和两个池化层提取熔池/匙孔图像的空域特征; 其次通过双向长短时记忆神经网络提取熔池/匙孔图像的时域特征; 最后将时空特征输入全连接网络进行特征降维并用 SOFTMAX 函数进行分类。所提方法建立了从熔池/匙孔图像到熔透类型的端到端的识别模型。试验结果表明, 所提方法在测试集上的识别准确率达 99.26%, 对单幅图像的识别时间为 9.43 ms, 可满足在线监测的需求。进一步的对比试验表明所提方法对学习率和优化器的变化具有鲁棒性。

**关键词** 激光技术; 模式识别; 熔透状态; 深度学习; 时序信息; 卷积神经网络; 长短时记忆神经网络

**中图分类号** TG403

**文献标志码** A

**doi:** 10.3788/CJL202148.0602119

### 1 引言

激光焊接具有深宽比大、效率高、热影响区和焊接变形小等优点<sup>[1]</sup>, 且激光焊接所需的能量可以得到精确的控制和优化, 因此激光焊接成为了自动化焊接的重要方式<sup>[2]</sup>。激光焊接技术在航空航天、汽车、电子等领域的精密零件焊接中得到了广泛的应用。然而, 由于激光焊接过程极其复杂, 等离子体、能量损失、焦点位置等因素均会影响熔透程度和最终成型, 对激光焊接熔透状态的控制一直是学术界的难点<sup>[3]</sup>。为了实现对熔透状态的精准控制, 首先要实现对熔透状态的精准识别。

在对熔透状态的识别中, 常用的监测信号主要有视觉信号<sup>[4-6]</sup>、声信号<sup>[7-8]</sup>、光谱信号<sup>[9]</sup>、温度信号<sup>[10]</sup>以及多源融合信号<sup>[11-12]</sup>。鉴于人类从自然界获取的信息大部分来自于视觉, 基于视觉信号的激光焊接熔透状态识别得到了最广泛的研究和应用<sup>[13]</sup>。但是激光焊接过程中存在等离子蒸气和飞

溅等干扰, 且相邻熔透状态之间差异极小, 这些问题给基于视觉的激光焊接熔透状态识别带来了极大的挑战。目前研究主要有基于特征工程<sup>[14-16]</sup>和基于深度学习<sup>[17-19]</sup>的思路来应对上述挑战。文献[14]首先对熔池/匙孔图像进行滤波和二值化预处理, 其次分别提取了熔池和匙孔的宽度和面积, 最后通过 BP 神经网络建立了特征与焊接状态之间的对应关系。文献[15]先对原始图像进行阈值分割、边缘提取等处理, 然后提取了熔池、匙孔、穿透孔、焊缝 4 方面特征, 再将得到的特征表示与标定的状态特征进行比较以判断熔透状态。文献[16]先对图像进行预处理, 然后进行人工设计并提取了蒸气图像的 15 种特征, 再采用顺序向前特征选择方法筛选最优的特征子集, 最后采用支持向量机(SVM)对特征进行分类。基于特征工程的思路主要遵循从预处理(如去噪、滤波、增强、边缘提取等)到特征(如长、宽、面积、周长、中心线等)提取再到特征分类(如感知机、支持向量机、阈值判断等)的技术路线。该思路可解释性

收稿日期: 2020-10-09; 修回日期: 2020-11-02; 录用日期: 2020-11-12

基金项目: 国家自然科学基金(51905091)、中央高校基本科研业务费专项资金、东华大学研究生创新基金(CUSF-DH-D-2020053)

\*E-mail: bao@dhu.edu.cn

强,但是存在中间子任务多的问题,该问题不但导致整体识别效率降低而且无法实现全局优化,更重要的是数据价值无法得到充分发挥。为了完整地描述熔池/匙孔的状态,文献[17]提出了将焊接过程的声信号转化为时频图以辅助判断,并将其与熔池/匙孔图像一起输入卷积神经网络(CNN)去自适应地提取熔透状态的特征。文献[18]针对部分熔透、中等熔透、全熔透、过度熔透 4 种状态的识别问题,研究了不同的图像输入尺寸及 CNN 的结构对识别性能的影响,并采用 ReLU 激活函数和 Dropout 方法防止模型过拟合。文献[19]基于残差网络提高了模型的训练效率,并采用 Center loss 促使网络学习多种焊接状态之间的鉴别性特征。基于深度学习的思路可以实现从原始图像到熔透状态的端到端的识别,该思路不但可以大大提高整体识别效率而且可以充分发挥数据价值,在智能化技术与制造业深度融合的大趋势下,基于深度学习的方法正成为当今的研究热点<sup>[20-21]</sup>。但是,基于深度学习的方法非常依赖数据,数据量过少会有过拟合的风险;相邻熔透状态间的界限比较模糊,这不利于需要清晰监督信息的深度学习模型去学习。

熟练焊接工人在实际焊接过程中往往考虑一段焊接过程中的渐变信息来做决策,这说明焊接质量的变化是一个渐变过程。因此,本文提出一种融入时序信息的激光焊接熔透状态识别方法。时序信息的引入不但可以增加深度学习模型对熔透状态的判断依据,而且可以间接地增加数据量。所提方法的总体框架由基于 CNN 的空域特征提取模块(SFEM)、基于双向长短时记忆(BiLSTM)的时域特

征提取模块(TDFEM)、基于全连接(FC)和 SOFTMAX 函数的分类模块(CM)组成。

## 2 方 法

### 2.1 总体框架

本文提出的融合时序信息的激光焊接熔透状态识别方法的总体框架如图 1 所示。总体思路是先利用 CNN 提取图像序列中每一帧图像的空域特征,然后将空域特征序列输入 BiLSTM 网络以挖掘时域内的上下文信息,最后将时空特征输入 CM 进行分类。其中, SFEM 的输入为 8 张  $64 \text{ pixel} \times 64 \text{ pixel}$  的灰度图。该图像序列经过 32 个  $5 \times 5$  的卷积核之后变为 8 个  $32 \times 64 \times 64$  的特征张量,特征张量经过窗口大小为  $2 \times 2$  的最大值池化(MP)变为 8 个  $32 \times 32 \times 32$  的特征张量。第二个卷积(Conv)层中的卷积核数目为 64 个,其他超参数与第一层一致。Conv 层中的移动步长设置为 1, MP 过程的移动步长设置为 2。为了保证 Conv 操作前后的空间尺寸一致,需要对 Conv 的输入进行外围补 0,因此在代码实现中将该参数设置为 SAME。图像序列经过两层 Conv 和 MP 之后变为 8 个  $64 \times 16 \times 16$  的特征张量。此时,将该 4 维张量转化为 2 维特征矩阵作为 TDFEM 的输入,该特征矩阵共有 8 个时间步,每个时间步的特征向量维度为  $64 \times 16 \times 16$ 。此外,将 BiLSTM 网络中的隐含单元设置为 64 个,最后将 TDFEM 的输出结果连接至 FC 层,输出结果经 SOFTMAX 函数得到最终的熔透状态识别结果。本文共涉及未熔透(LP)、最佳熔透(OP)和过度熔透(EP)三种类型。

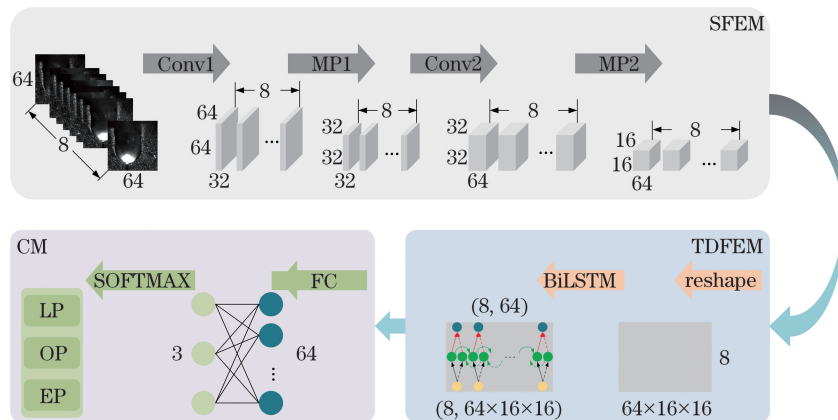


图 1 总体框架

Fig. 1 Overall framework

### 2.2 空域特征提取

由图 1 可以看出,本文算法主要分为两个核心

阶段:基于 CNN 的空域特征提取阶段和基于 BiLSTM 的时域特征提取阶段。在空域特征提取阶

段,图像信号的前向传播过程如下:假设第  $l_c$  层为 Conv 层,第  $l_c-1$  层为池化层或输入层。则第  $l_c$  层的计算公式为

$$x_{j_c}^{l_c} = f_c \left( \sum_{i_c=1}^{n_c} x_{i_c}^{l_c-1} \times k_{j_c}^{l_c} + b_{j_c}^{l_c} \right), \quad (1)$$

式中:  $x_{j_c}^{l_c}$  为第  $l_c$  层的第  $j_c$  个特征图;  $f_c$  为 Conv 层的激活函数;  $x_{i_c}^{l_c-1}$  为  $l_c-1$  层的一个输入特征图;  $n_c$  为第  $l_c-1$  层的特征图的数量;  $b_{j_c}^{l_c}$  为卷积层的偏置项;  $k_{j_c}^{l_c}$  为卷积核即共享权重。(1)式右边表示的是用  $l_c$  层的第  $j_c$  个卷积核与第  $l_c-1$  层所有的特征图  $x_{i_c}^{l_c-1}$  进行卷积运算并求和,然后加上一个偏置项,再将所得结果输入激活函数,得到输出。

假设第  $l_p$  层为池化层(也称为下采样层),第  $l_p-1$  层为 Conv 层。则第  $l_p$  层的计算公式为

$$x_{j_p}^{l_p} = f_p [\beta \times \text{down}(x_{j_p}^{l_p-1}) + b_{j_p}^{l_p}], \quad (2)$$

式中:  $x_{j_p}^{l_p-1}$  为  $l_p-1$  层的第  $j_p$  个特征图;  $x_{j_p}^{l_p}$  为第  $l_p$  层的第  $j_p$  个特征图;  $f_p$  为池化层的激活函数; down 为下采样函数;  $\beta$  为下采样系数;  $b_{j_p}^{l_p}$  为池化层的偏置项。

### 2.3 时域特征提取

本文采用图 2 所示的 BiLSTM 网络来提取熔池/匙孔图像序列内的时序特征。BiLSTM 由正向 LSTM(F-LSTM)和反向 LSTM(R-LSTM)组成。设  $t$  时刻的输入为  $x_t$ ,则 BiLSTM 的内涵是对  $t$  时刻的判断不仅依赖于  $t$  时刻之前的信息还依赖于  $t$  时刻之后的信息。对时间上下文的关联性的挖掘可以提高模型识别当前时刻的熔透状态的可靠性。其中利用 F-LSTM 对当前序列进行正向递归以挖掘  $x_1, \dots, x_{t-1}$  对  $x_t$  的影响,在 F-LSTM 的内部迭代中以  $x_1$  作为第一个时刻的输入;利用 R-LSTM 对当前序列进行反向递归以挖掘  $x_8, \dots, x_{t+1}$  对  $x_t$  的影响,在 R-LSTM 网络的内部迭代中以  $x_8$  作为第一个时刻的输入。在网络训练过程中,熔池/匙孔图像序列被同时输入 F-LSTM 和 R-LSTM 以获得正向输出  $\vec{h}_t$  和反向输出  $\overleftarrow{h}_t$ ,然后将相同输入情况下的正向输出  $\vec{h}_t$  和反向输出  $\overleftarrow{h}_t$  相加,得到该输入的最终输出  $H_t$ 。根据图 1,本文涉及的 BiLSTM 在每一个时间步的输入是长度为  $64 \times 16 \times 16$  的一维特征向量,共需 8 个时间步将当前序列递归完毕,并将最后一个时间步的输出  $H_8$  作为该模块的最终输出。

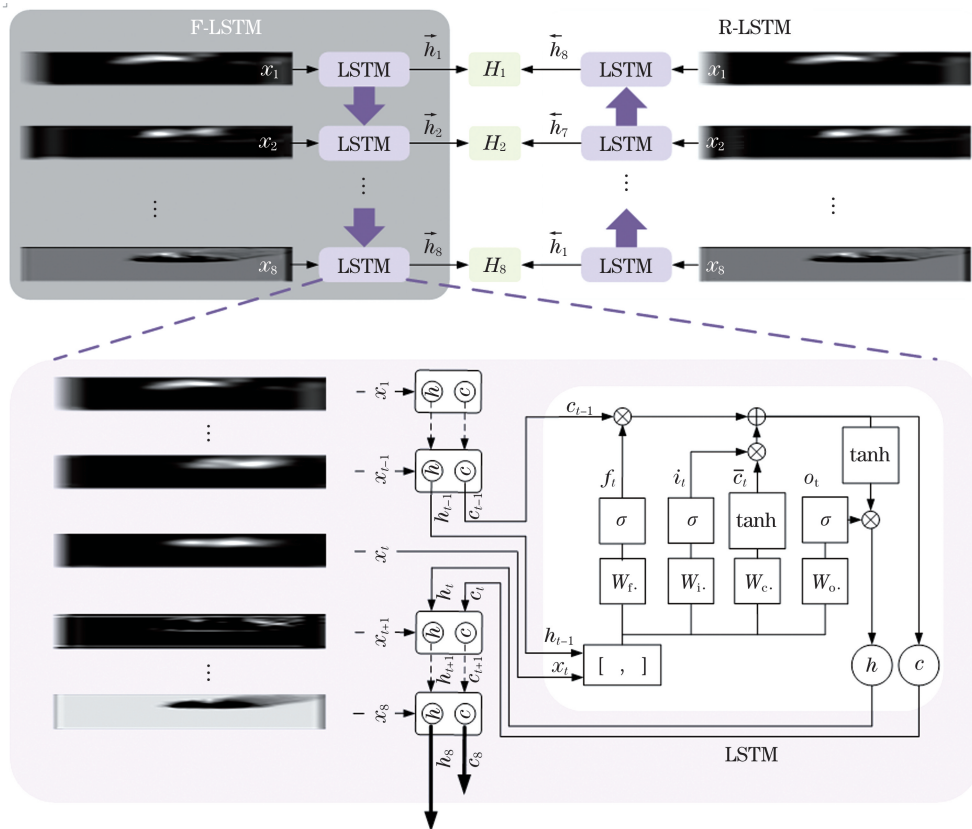


图 2 BiLSTM 网络结构

Fig. 2 Structure of BiLSTM network



由图 2 可知, BiLSTM 的基础构成是 LSTM 网络。LSTM 网络由三个门限结构来控制保存长期记忆的单元的状态。图 2 中,  $c$  用于存储记忆信息, 就长期记忆和短期记忆而言, LSTM 由  $c_t$  保存长期记忆,  $\bar{c}_t$  保存短期记忆。在三个门限结构中, 遗忘门  $f_t$  决定了历史信息有多少可以保留在长期记忆  $c_t$  中; 输入门  $i_t$  决定当前输入网络的信息有多少可以加入长期记忆  $c_t$  中; 输出门  $o_t$  控制的是汇总后的长短期记忆有多少可以作为当前输出的信息。  $h$  用于表示网络的输出,  $h_t$  则是当前时刻的输出。用 “ $\cdot$ ” 表示矩阵乘法, “ $\circ$ ” 表示相同位置的元素相乘, 则 LSTM 产生信息记忆功能的表达式为

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f), \quad (3)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \quad (4)$$

$$\bar{c}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c), \quad (5)$$

$$c_t = f_t \circ c_{t-1} + i_t \circ \bar{c}_t, \quad (6)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \quad (7)$$

$$h_t = o_t \circ \tanh(c_t), \quad (8)$$

式中:  $W_f$  和  $b_f$  为遗忘门的权重矩阵及偏置项;  $W_i$  和  $b_i$  为输入门的权重矩阵及偏置项;  $W_c$  和  $b_c$  为由当前输入生成短期记忆  $\bar{c}_t$  的权重矩阵和偏置项;  $W_o$  和  $b_o$  为输出门的权重矩阵及偏置项;  $\sigma$  和  $\tanh$  为两种不同的激活函数, 二者的表达式分别为

$$\sigma(x) = \frac{1}{1 + \exp(-x)}, \quad (9)$$

$$\tanh(x) = \frac{\exp(x) - \exp(-x)}{\exp(x) + \exp(-x)}. \quad (10)$$

根据(9)和(10)式, 易知  $\sigma$  函数的值域为  $(0, 1)$ ,  $\tanh$  函数的值域为  $(-1, 1)$ 。因此, 当  $\sigma$  函数的输入值趋近于 0 时, 意味着信息都不能通过, 当  $\sigma$  函数的输入值趋近于 1 时, 意味着信息都能通过。在生成短期记忆  $\bar{c}_t$  时, 采用  $\tanh$  函数是为了让短期记

忆以 0 为中心,  $\tanh$  函数在 0 附近的梯度较大, 可以加速模型收敛。

### 2.4 时空特征分类

CM 的作用是将前述模块得到的时空特征映射到具体的熔透状态上。CM 首先通过 FC 层将时空特征降维到与类别数一致, 其次通过 SOFTMAX 函数将不同类别对应的特征值转化为类别概率。用  $x_{in}$  表示全连接层的输入,  $x_{out}$  表示全连接层的输出,  $f_{fc}$  表示全连接层的激活函数,  $W_{fc}$  表示全连接层的权重矩阵,  $b_{fc}$  表示全连接层的偏置, 则全连接层的输出为

$$x_{out} = f_{fc}(W_{fc} \times x_{in} + b_{fc}). \quad (11)$$

最终分类采用的 SOFTMAX 函数的表达式为

$$\text{SOFTMAX}(i) = \frac{\exp(x_{out, i})}{\sum_j \exp(x_{out, j})}, \quad (12)$$

式中:  $x_{out, i}$  表示全连接层的第  $i$  个输出,  $j$  表示全连接层输出神经元的个数。

## 3 试验与分析

### 3.1 激光焊接试验

本文进行焊接试验的系统如图 3 所示, 包括焊接模块和视觉模块两大部分。试验采用高速相机进行旁轴拍摄以观察熔池/匙孔区域。此外, 所选激光器的型号为 IPG-YLS6000, 辅助光源的波长为 808 nm, 滤光片的中心为 808 nm, 焊接母材选用长宽高分别为 300, 120, 2 mm 的 SS304 不锈钢板, 保护气采用流量为  $25 \text{ L} \cdot \text{min}^{-1}$  的氩气, 焊接速度为  $2 \text{ m} \cdot \text{min}^{-1}$ , 离焦量为 +8 mm, 相机型号为 CP70-2M。

通过改变激光功率的大小可以获得不同的熔透类型。经过大量试焊, 发现当激光功率为 1800 W 时可以获得 OP 状态。因此, 本文在激光功率为 1800 W 的基础上分别增加和减小 50 W 作为获取

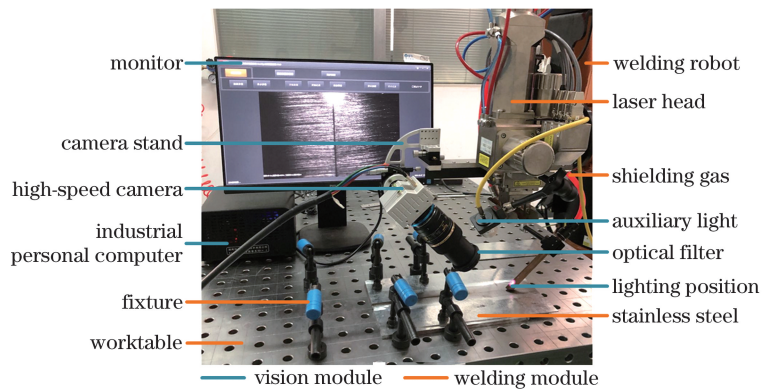


图 3 试验系统

Fig. 3 Test system



OP 状态的工艺参数;然后在 1750 W 的基础上逐步递减 100 W 以获得 LP 状态,在 1850 W 的基础上逐步递增 100 W 以获得 EP 状态。在焊接完成后,不同熔透状态对应的焊缝背面和熔池/匙孔图像如图 4 所示。随着熔透程度的增加,熔池区域的宽度会有小幅度的增加<sup>[22]</sup>,匙孔区域的等离子蒸气和飞

溅会越来越多且规则性会越来越差<sup>[23]</sup>。在特征工程的思路下,这种图像特点要求大量的预处理和精细的特征设计工作,然而基于深度学习的思路则可以以端到端的方式建立熔池/匙孔图像与熔透状态之间的关系。所获数据集的分布和划分情况如表 1 所示。

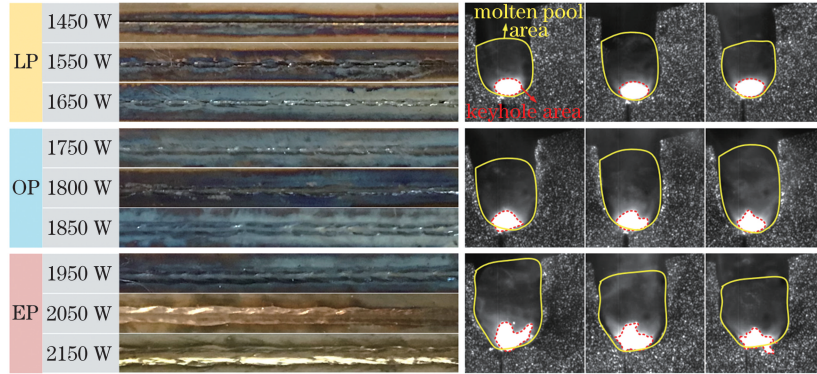


图 4 不同熔透状态对应的焊缝背面和熔池/匙孔图像

Fig. 4 Images of backside of weld joint and molten pool/keyhole corresponding to different penetration states

表 1 数据分布和数据集划分情况

Table 1 Distribution and division of datasets

| Dataset | LP    | OP    | EP    | Total |
|---------|-------|-------|-------|-------|
| Train   | 12817 | 12807 | 12976 | 38600 |
| Valid   | 3653  | 3641  | 3736  | 11030 |
| Test    | 1802  | 1872  | 1848  | 5522  |
| Total   | 18272 | 18320 | 18560 | 55152 |

### 3.2 算法仿真

#### 3.2.1 仿真设置

在算法的优越性测试方面:本文将学习率(LR)设置为  $10^{-3}$ ,损失函数采用交叉熵,优化器(OM)选用带动量的随机梯度下降法(SGD-M),动量因子设为 0.9。此外,将原始数据集按照表 1 划分为训练集、验证集、测试集三部分。训练过程中将批次设置为 8 以加速训练。为了模拟实际工业环境,本文算法在普通的 Windows 10 操作系统、NVIDIA 1060 显卡、3 GB 显存的硬件环境以及 Pytorch 深度学习框架下进行算法性能仿真。为了体现所提方法的优越性,将所提方法与 CNN 模型(与本文 CNN 结构相同)、LSTM 模型(与本文 LSTM 结构相同)、CNN-LSTM 模型(仅单向 LSTM)进行对比。

在算法的鲁棒性测试方面,鉴于深度学习领域一般都选择交叉熵作为损失函数,或者是根据特定任务针对性地设计损失函数,而 LR 和 OM 的选择对模型的性能影响较大,因此本文将在保证其他参

数固定的情况下,采用不同的 LR 和 OM 来验证模型对二者的鲁棒性。

在评价指标方面,采用准确率( $A_{ACC}$ )、精确率( $P_{PRE}$ )、召回率( $R_{REC}$ )、 $F_{F1-score}$  4 个指标对相关方法进行综合评估。各指标的定义为

$$A_{ACC} = \frac{T_P + T_N}{T_P + T_N + F_P + F_N}, \quad (13)$$

$$P_{PRE} = \frac{T_P}{T_P + F_P}, \quad (14)$$

$$R_{REC} = \frac{T_P}{T_P + F_N}, \quad (15)$$

$$F_{F1-score} = \frac{2 \times P_{PRE} \times R_{REC}}{P_{PRE} + R_{REC}}, \quad (16)$$

式中:  $T_P$  为实际为正类同时被识别为正类的样本数量;  $T_N$  为实际为负类同时被识别为负类的样本数量;  $F_P$  为实际为负类同时被识别为正类的样本数量;  $F_N$  为实际为正类同时被识别为负类的样本数量;  $A_{ACC}$  为所有预测正确的样本占总样本的比例;  $P_{PRE}$  为实际为正类同时被识别为正类的样本数量

占有所有被识别为正类的样本总数的比例;  $R_{REC}$  为实际为正类同时被识别为正类的样本数量占被预测为正类的样本数量的比例,  $F_{F1-score}$  为  $P_{PRE}$  和  $R_{REC}$  的调和平均值。

### 3.2.2 优越性分析

不同方法的训练过程如图 5 所示。从图中可以看出:融合了时序信息的 CNN-LSTM 和 CNN-BiLSTM 方法的准确率和收敛速度均大于未加入时序信息的 CNN 方法。其中, CNN-BiLSTM 方法的  $A_{ACC}$  经过 5000 次迭代即可收敛于 1 且收敛过程比较稳定, 而 CNN 方法的  $A_{ACC}$  则只能收敛于 0.93 左右。

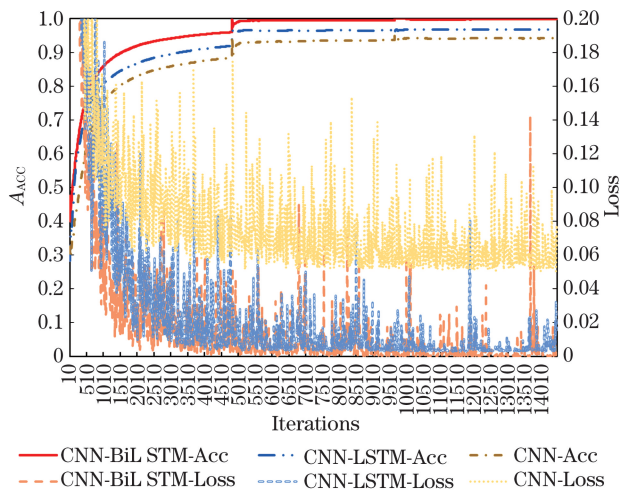


图 5 不同方法的训练过程对比

Fig. 5 Comparison of training processes of different methods

在训练过程中, 当对所有训练数据完成一次遍历(epoch)后, 用验证集对 CNN-BiLSTM 算法的性

能进行验证, 不同 epoch 后模型在验证集上的表现如图 6 所示。从图中可以看出, CNN-BiLSTM 算法在训练和验证集上的差距不大, 这说明所提方法没有发生明显的过拟合现象。验证精度之所以略大于训练精度是因为训练精度采用的是当前 Epoch 周期内所有批次结果的平均。

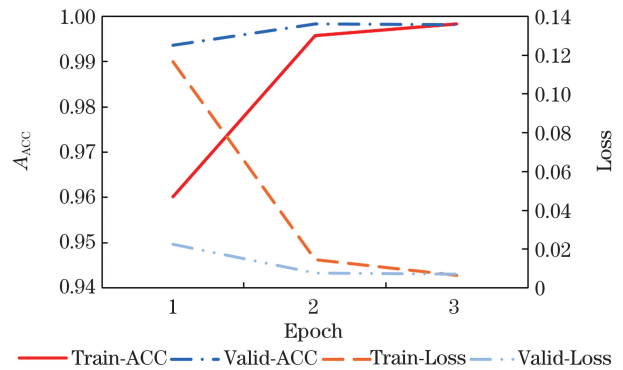


图 6 CNN-BiLSTM 方法在训练集和验证集上的对比  
Fig. 6 Comparison of CNN-BiLSTM method on training set and validation set

模型训练完成后, 将其保存为可测试文件并在测试集上进行测试。不同方法在测试集上的各项指标如表 2 所示。从表 2 中可以看出, 所提 CNN-BiLSTM 方法在测试集上的总体识别准确率  $A_{ACC}$  和综合评价指标  $F_{F1-score}$  均可达 99.26%, 这远高于 CNN 方法。虽然 CNN-LSTM 方法的各项指标也可达到 97% 左右, 但是 CNN-LSTM 方法仅考虑了当前输入时刻在时域上的上文信息, 并没有考虑当前输入在时域上的下文信息, 因此 CNN-LSTM 方法在精度方面的指标仍低于 CNN-BiLSTM 方法。

表 2 不同方法在测试集上的测试结果

Table 2 Test results of different methods on test set

| Parameter          | CNN   | CNN-LSTM | CNN-BiLSTM |
|--------------------|-------|----------|------------|
| $A_{ACC}/\%$       | 92.99 | 96.99    | 99.26      |
| $P_{PRE}/\%$       | 92.99 | 97.00    | 99.25      |
| $R_{REC}/\%$       | 93.03 | 97.01    | 99.26      |
| $F_{F1-score}/\%$  | 93.01 | 97.01    | 99.26      |
| Model size /MB     | 4.30  | 16.66    | 33.11      |
| Training time /min | 3.15  | 7.93     | 8.68       |
| Test time /ms      | 1.76  | 9.41     | 9.43       |

鉴于当今的工业计算机一般为 16G 内存和 1 TB 硬盘容量以上, 且工业相机的帧率一般不超过 100 Hz。因此, 虽然时序信息的引入以及 LSTM 网络携带的参数会导致时序模型的体积相对较大, 训练

时间、测试时间均相对较长, 但是由于计算机水平的发展, 该结果尚不会对熔池状态监测过程造成影响。

图 7 展示了不同方法在测试集上的分类结果的混淆矩阵。从混淆矩阵中可以发现, 当 CNN 和

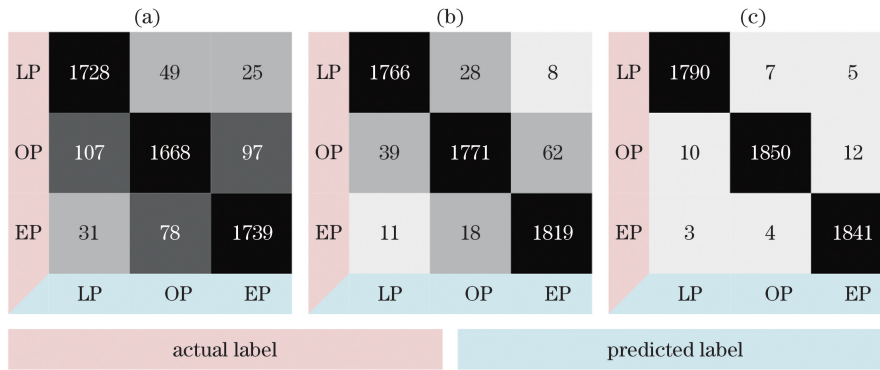


图 7 不同方法分类结果的混淆矩阵。(a) CNN; (b) CNN-LSTM; (c) CNN-BiLSTM

Fig. 7 Confusion matrices of classification results of different methods. (a) CNN; (b) CNN-LSTM; (c) CNN-BiLSTM

CNN-LSTM 方法对熔透状态错误分类时,易将熔透状态分类为与真实标签相近的那一类,未熔透状态则易被错分为熔透状态。所提方法对易错分的情况有一定的改善。

### 3.2.3 鲁棒性分析

为了验证所提方法的鲁棒性,本文在其他参数保持不变的情况下将 LR 分别设置为  $10^{-2}$ 、 $10^{-4}$ 、 $10^{-5}$ ,将 OM 分别设置为 Adagrad、Adadelata、Adam 以对 CNN-BiLSTM 算法进行测试。不同超参数下的训练过程如图 8 所示。

从图 8 可以看出:当 LR 与本文设置的  $10^{-3}$  接

近时,不同超参数对应的训练过程的收敛情况较为一致,最终收敛的值也比较接近;当 LR 为  $10^{-5}$  时,模型并没有收敛,这说明 LR 太小的情况下模型发生了欠拟合现象。当 OM 设置为 Adagrad 和 Adam 时,训练过程的收敛情况与本文采用的 SGD-M 优化器类似;当 OM 选为 Adadelata 时,模型对训练集遍历三次后无法完成收敛,这主要是因为 Adadelata 在训练中后期容易陷入局部最优。训练完成后,将模型保存为可测试文件并在测试集上进行测试,测试结果如表 3 所示。

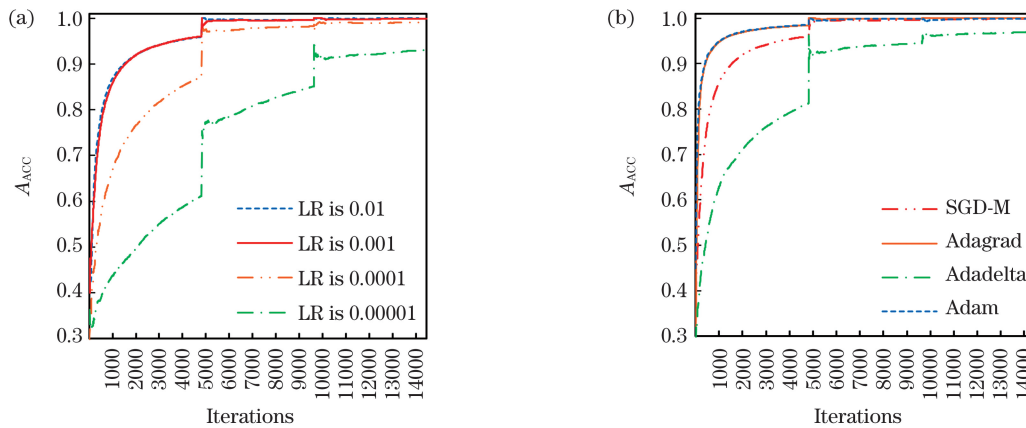


图 8 不同超参数对应的训练过程。(a)不同学习率;(b)不同优化器

Fig. 8 Training processes corresponding to different hyperparameters. (a) Different learning rates; (b) different optimizers

表 3 不同超参数对应的测试结果

Table 3 Test results for different hyperparameters

| Parameter           | LR        |           |           | OM        |         |       |
|---------------------|-----------|-----------|-----------|-----------|---------|-------|
|                     | $10^{-5}$ | $10^{-4}$ | $10^{-2}$ | Adadelata | Adagrad | Adam  |
| $A_{ACC} / \%$      | 93.14     | 99.11     | 99.35     | 97.66     | 99.18   | 99.31 |
| $P_{PRE} / \%$      | 93.70     | 99.11     | 99.35     | 97.67     | 99.19   | 99.31 |
| $R_{REC} / \%$      | 93.14     | 99.11     | 99.35     | 97.67     | 99.18   | 99.31 |
| $F_{FI-score} / \%$ | 93.14     | 99.11     | 99.35     | 97.66     | 99.19   | 99.31 |



由于模型的结构和参数量没有改变,因此模型的大小、训练时间、测试时间与前文一致。在精度指标方面:当 LR 为  $10^{-5}$  时模型的精度最低,这主要是因为过低的 LR 导致模型处于有待训练的欠拟合状态。然而,当 LR 在本文设置的  $10^{-3}$  附近或者更换 OM 时,网络的各项精度指标均可保持在 97.66% 以上,这说明所提方法对 LR 和 OM 的变化具有一定的鲁棒性。

## 4 结 论

为了解决不同熔透状态间鉴别性低且熔池区域存在等离子蒸气及飞溅干扰的问题,在基于深度学习的激光焊接熔透状态识别过程中考虑了当前输入在时域上的上下文信息,该方法不但可以增加对当前输入的判断依据还可以间接地起到增加数据集的作用。所提方法经过 5000 次迭代即可收敛于 1 且训练过程稳定;所提方法在验证集上没有发生过拟合现象;所提方法在测试集上的总体识别准确率达 99.26%;所提方法在普通计算机上识别单幅图像仅需 9.43 ms;所提方法对学习率和优化器的变化具有鲁棒性。因此,所提方法满足基于视觉的激光焊接熔透状态监测的实际需求。

虽然所提方法展现了很多优势,但是在今后的研究中仍需关注以下方面。在研究对象方面:熔透状态除了本文设定的三种之外,还有很多中间过渡状态,如部分熔透和完全熔透等。因此,对熔透状态进行更精细的划分将是以后的一个关注点。在模型结构方面:本文用于提取时序信息的骨架网络是经典的 LSTM 网络,该网络的一大缺点就是在时间维度上存在递归操作,这会导致网络不能并行运算以至于耗费更多的训练和测试时间。因此,如何从结构上解放网络的并行运算能力将是未来的一个研究点。在模型优化方面:时序信息的引入会不可避免地增大模型的体积以及增加模型的训练和测试时间。在智能制造的浪潮下,在焊接边缘节点布置的传感器将越来越多,能分配给视觉模块的资源将更有限。因此,如何在不显著降低准确性的前提下实现时序模型的轻量化将是未来的一个研究点。

## 参 考 文 献

- [1] Lu W, Ma X Y, Gong S L, et al. Effect of laser welding process on weld formation and mechanical property of GH188 alloy [J]. Laser & Optoelectronics Progress, 2017, 54(4): 041419.
- [2] Chen Z Q, Gao X D, Wang Y, et al. Weldment back of weld width prediction based on neural network during high-power laser welding [J]. Transactions of the China Welding Institution, 2018, 39(11): 48-52, 131.  
陈子琴, 高向东, 王煜, 等. 大功率激光焊背面焊缝宽度神经网络预测 [J]. 焊接学报, 2018, 39(11): 48-52, 131.
- [3] Gao X D, Liang J B, Liu G Q, et al. Identification of high-power fiber laser welding penetration based on fuzzy clustering algorithm [J]. Transactions of the China Welding Institution, 2017, 38(5): 22-25, 129-130.  
高向东, 梁剑斌, 刘桂谦, 等. 大功率光纤激光焊熔透状态模糊聚类识别方法 [J]. 焊接学报, 2017, 38(5): 22-25, 129-130.
- [4] Song H Y, Chen J T, Song J, et al. On-line inspection of laser welding quality based on infrared thermal imaging technology [J]. Laser Journal, 2020, 41(8): 45-49.  
宋海燕, 陈继涛, 宋娟, 等. 基于红外热成像技术的激光焊接质量在线检测 [J]. 激光杂志, 2020, 41(8): 45-49.
- [5] Li S R, An S H, Wang C M, et al. Effect of shape and behavior of high power laser welding keyhole on weld morphology and mechanical properties [J]. Applied Laser, 2019, 39(6): 956-960.  
李尚仁, 安升辉, 王春明, 等. 高功率激光焊接匙孔形态行为对焊缝成形及力学性能的影响 [J]. 应用激光, 2019, 39(6): 956-960.
- [6] Zhang Y, Li F Z, Liang Z C, et al. Correlation analysis of penetration based on keyhole and plasma plume in laser welding [J]. Journal of Materials Processing Technology, 2018, 256: 1-12.
- [7] Yusof M F M, Ishak M, Ghazali M F. Classification of weld penetration condition through synchrosqueezed-wavelet analysis of sound signal acquired from pulse mode laser welding process [J]. Journal of Materials Processing Technology, 2020, 279: 116559.
- [8] Ao S S, Luo Z, Huang Z Y, et al. Numerical calculation and experimental analysis based on acoustic emission of molten pool keyhole in laser deep penetration welding [J]. Transactions of the China Welding Institution, 2016, 37(11): 93-98, 133.  
敖三三, 罗震, 黄尊月, 等. 激光深熔焊过程熔池小孔发声数值计算与试验分析 [J]. 焊接学报, 2016, 37(11): 93-98, 133.

- [9] Li Z M, Gao X D, Zhang N F. Analysis of high-power disk laser welding status based on multi-feature fusion[J]. *Laser Technology*, 2017, 41(5): 764-768.  
李竹曼, 高向东, 张南峰. 大功率碟片激光焊接状态多特征融合分析法[J]. *激光技术*, 2017, 41(5): 764-768.
- [10] Chen G Y, Zhang M J, Zhao Z, et al. Measurements of laser-induced plasma temperature field in deep penetration laser welding [J]. *Optics & Laser Technology*, 2013, 45: 551-557.
- [11] Wang C M, Wu S P, Hu L J, et al. Identification of different laser welding penetration states based on multi-sensor fusion [J]. *Chinese Journal of Lasers*, 2007, 34(4): 538-542.  
王春明, 吴松坪, 胡伦骥, 等. 基于多传感器融合的激光焊接熔透状态的识别[J]. *中国激光*, 2007, 34(4): 538-542.
- [12] Zhang Y X, You D Y, Gao X D, et al. Online monitoring of welding status based on a DBN model during laser welding [J]. *Engineering*, 2019, 5(4): 671-678.
- [13] Gu J, Zhang L L, Wang J C. Application of machine vision in automatic welding [J]. *Applied Laser*, 2018, 38(6): 1038-1043.  
顾俊, 张玲玲, 王健超. 机器视觉在自动化焊接中的应用[J]. *应用激光*, 2018, 38(6): 1038-1043.
- [14] Gong J F, Li H Z, Li L Q, et al. Quality monitoring technology of laser welding process based on coaxial image sensing [J]. *Transactions of the China Welding Institution*, 2019, 40(1): 37-42, 162-163.  
宫建锋, 李慧知, 李俐群, 等. 基于同轴图像传感的激光焊接过程质量监测技术[J]. *焊接学报*, 2019, 40(1): 37-42, 162-163.
- [15] Zhang Y, Liu T W, Li B, et al. Simultaneous monitoring of penetration status and joint tracking during laser keyhole welding [J]. *IEEE/ASME Transactions on Mechatronics*, 2019, 24(4): 1732-1742.
- [16] Wang T, Chen J Q, Gao X D, et al. Real-time monitoring for disk laser welding based on feature selection and SVM [J]. *Applied Sciences*, 2017, 7(9): 884.
- [17] Wu D, Huang Y M, Zhang P L, et al. Visual-acoustic penetration recognition in variable polarity plasma arc welding process using hybrid deep learning approach [J]. *IEEE Access*, 2020, 8: 120417-120428.
- [18] Zhang Z H, Li B, Zhang W F, et al. Real-time penetration state monitoring using convolutional neural network for laser welding of tailor rolled blanks [J]. *Journal of Manufacturing Systems*, 2020, 54: 348-360.
- [19] Xia C Y, Pan Z X, Fei Z Y, et al. Vision based defects detection for keyhole TIG welding using deep learning with visual explanation [J]. *Journal of Manufacturing Processes*, 2020, 56: 845-855.
- [20] Cai W, Wang J Z, Jiang P, et al. Application of sensing techniques and artificial intelligence-based methods to laser welding real-time monitoring: a critical review of recent literature [J]. *Journal of Manufacturing Systems*, 2020, 57: 1-18.
- [21] Wang B C, Hu S J, Sun L, et al. Intelligent welding system technologies: state-of-the-art review and perspectives [J]. *Journal of Manufacturing Systems*, 2020, 56: 373-391.
- [22] Zhao L, Cao Z, Zou J L, et al. Morphology characteristics of keyhole in high power fiber laser deep penetration welding [J]. *Chinese Journal of Lasers*, 2020, 47(11): 1102005.  
赵乐, 曹政, 邹江林, 等. 高功率光纤激光深熔焊接小孔的形貌特征[J]. *中国激光*, 2020, 47(11): 1102005.
- [23] Han X, Zhao Y, Zou J L, et al. Analysis of plume formation reasons in laser deep penetration welding based on visual observation [J]. *Chinese Journal of Lasers*, 2020, 47(6): 0602004.  
韩雪, 赵宇, 邹江林, 等. 基于可视化观察的光纤激光深熔焊接羽辉形成原因分析[J]. *中国激光*, 2020, 47(6): 0602004.

# Laser Welding Penetration State Recognition Method Fused with Timing Information

Liu Tianyuan<sup>1</sup>, Bao Jinsong<sup>1\*</sup>, Wang Junliang<sup>1</sup>, Gu Jun<sup>2</sup>

<sup>1</sup>*Institute of Intelligent Manufacturing, College of Mechanical Engineering, Donghua University, Shanghai 201600, China;*

<sup>2</sup>*Shanghai Institute of Laser Technology, Shanghai 200235, China*

## Abstract

**Objective** It is essential to initially establish a precise recognition model to achieve accurate control for a penetration state in laser welding. Although the recognition method of the penetration state using visual signals is widely proposed, there are still interferences, such as plasma, vapor, and spattering, in the laser welding process. Besides, there is no significant change in adjacent penetration state. These are the problems in vision-based recognition of the laser welding penetration state. Feature engineering and deep learning seem to be the only methods to solve these problems. Although the feature engineering-based method is interpretable, it requires many subtasks to decrease overall recognition efficiency. Also, the value of data cannot be fully developed. On the other hand, the deep learning-based method realizes an end-to-end recognition from the original image to the penetration state. It improves the overall recognition efficiency and data value. The deep learning-based method becomes a major research focus because of integrating intelligent technologies into manufacturing systems. However, deep learning-based methods require a large amount of data, because fewer data result in overfitting. The boundary between adjacent penetration states is unclear, making it difficult for supervised learning methods to be applied. Inspired by the fact that skilled welders consider asymptotic information in deciding on the welding process, we propose a laser welding penetration state identification method that incorporates timing information. The timing information is expected to improve the determining factor of a deep learning method for the weld penetration state and increase the amount of data.

**Methods** The frame of the proposed method consists of a spatial feature extraction module (SFEM) using a convolutional neural network (CNN), a time domain feature extraction module (TDFEM) using a bi-directional long short-term memory (BiLSTM) neural network, and a classification module (CM) using a SOFTMAX function. In the SFEM, we used two convolutional layers and two max-pooling layers to extract the input image sequence's spatial features. Afterward, we applied TDFEM to extract the features in the time domain from the input sequence. In the TDFEM, the feature sequence was simultaneously input into the forward LSTM and reversed the LSTM to obtain the forward and reverse outputs. Then, we summed the forward and reverse outputs of the same input as the final output of the current input. In the CM, we first input the spatiotemporal features into the fully connected network for dimensionality reduction. Subsequently, we mapped the low-dimension features to categorical probabilities using the SOFTMAX function. For data acquisition, the optimum penetration condition was obtained through a welding test. We incremented the laser power corresponding to the optimum penetration state, and then decremented it to obtain excessive penetration and incomplete penetration conditions.

**Results and Discussions** Figure 5 shows that the CNN-BiLSTM method's accuracy converges to 1 after 5000 iterations, whereas the CNN method's accuracy only converges to approximately 0.93. Using the CNN-BiLSTM method, the difference in accuracy on training and validation sets is insignificant, suggesting no overfitting using the proposed method (Fig. 6). The identification accuracy and overall evaluation index of the CNN-BiLSTM method reach 99.26% on the test set, much higher than the those of the conventional CNN method. Although various CNN-LSTM method indicators are about 97%, the CNN-LSTM method only considers the previous information of current input in the time domain without considering subsequent information of current input in the time domain. The proposed method takes only 9.43 ms to identify a single image in a PC (Table 2). The CNN and CNN-LSTM methods misclassify the penetration state as the true label. Moreover, Fig. 7 shows that the proposed method can suppress misclassification. In this paper, the training process's convergence tends to be consistent when the learning rate (LR) is close to  $10^{-3}$ . The model does not converge when the LR is  $10^{-5}$ , suggesting no overfitting. When the optimizer (OM) is set to Adagrad or Adam, the training process's convergence is similar to that of the stochastic



gradient descent method applicable to this paper. The proposed method could not converge within three epochs when the Adadelta OM is used because the Adadelta easily falls into the local optimum in the middle and later training stages (Fig. 8). Table 3 shows that all accuracy metrics of CNN-BiLSTM are above 97.66% when the LR is around  $10^{-3}$  or OM is replaced, which suggests that the proposed method is robust.

**Conclusions** In this paper, timing information was not considered in deep learning-based methods for penetration state recognition. Our proposed method, CNN-BiLSTM, can adaptively extract spatiotemporal context information, as the method demonstrated good convergences and stability. The introduction of temporal information can indirectly play the role in data augmentation, and the proposed method does not overfit. The overall recognition accuracy of the proposed method on the test set is 99.26%. As such, the proposed method meets the standard requirements of vision-based laser welding in terms of penetration condition monitoring. Furthermore, the method is robust to changes in LR and optimizer. Although the proposed method has many advantages, the following aspects still need to be focused in future research. Better classification of penetration states will be a future interest choice in terms of research objects. Also, making the network structure fit the parallel computing system will be a good future research direction. In terms of model optimization, making the timing model lightweight without reducing the accuracy will be another good option.

**Key words** laser technique; pattern recognition; penetration state; deep learning; timing information; convolutional neural networks; long short-term memory neural networks

**OCIS codes** 140.3040; 140.5495; 100.4996